Joffe

Physics library
QC
1
.P5812
v.56
no.1

THE PHYSICAL REVIEW

VOLUME 56 *Second Series* NUMBER 1

JULY 1, 1939

Published for the

AMERICAN PHYSICAL SOCIETY

by the

AMERICAN INSTITUTE OF PHYSICS
Incorporated

LANCASTER, PA., AND NEW YORK, N. Y.

Physics

UNIVERSITY OF MICHIGAN LIBRARIES

THE PHYSICAL REVIEW

A journal of experimental and theoretical physics established by E. L. Nichols in 1893

Published by the American Institute of Physics for the

American Physical Society

OFFICERS

President

JOHN T. TATE
University of Minnesota
Minneapolis, Minnesota

Vice President

JOHN ZELENY
Yale University
New Haven, Connecticut

Secretary

W. L. SEVERINGHAUS
Columbia University
New York, New York

Treasurer

G. B. PEGRAM
Columbia University
New York, New York

Local Secretary for the Pacific Coast

PAUL KIRKPATRICK
Stanford University
California

Editor

JOHN T. TATE
University of Minnesota
Minneapolis, Minnesota

BOARD OF EDITORS

JOHN T. TATE, *Editor* J. W. BUCHTA, *Assistant Editor*

Associate Editors

K. T. BAINBRIDGE	A. L. HUGHES	W. BLEAKNEY
H. P. ROBERTSON	T. H. JOHNSON	G. BREIT
H. B. WAHLIN	F. A. JENKINS	D. M. DENNISON

Manuscripts for publication should be submitted to John T. Tate, University of Minnesota, Minneapolis, Minnesota.

Proof and all correspondence concerning papers in the process of publication, should be addressed to the Publications Manager, American Institute of Physics, 175 Fifth Avenue, New York, New York.

Subscription Price

United States and Possessions and Canada $15.00
Foreign $16.50

Back Numbers

Complete set: Series 2,[1] Vol. 1, 1913–Vol. 54, 1938—$390.00
Yearly rate: $16.50
Single copies: $1.50 each prior to July, 1929; $.75 each after July, 1929.

Subscriptions and orders for back numbers should be addressed to Prince and Lemon Streets, Lancaster, Pennsylvania, or to the American Institute of Physics, 175 Fifth Avenue, New York, New York.

Advertising is not accepted for the *Physical Review* but attention is called to the *Review of Scientific Instruments* and the *Journal of Applied Physics*, companion publications of the Institute, and media through which the field of physics can be reached most effectively. Rates and other information will be furnished on request to the Advertising Manager.

Changes of address and complaints of failure to receive the *Physical Review* in the case of members of the American Physical Society, should be addressed to the Treasurer; in the case of other subscribers, to the Publications Manager. New copies can be sent free in response to complaints of non-delivery only if notice of non-delivery is received within three months of date of issue.

[1] For Series 1 inquire of Physics Department, Cornell University, Ithaca, N. Y.

The Physical Review is published semi-monthly at Prince and Lemon Streets, Lancaster, Pennsylvania.

Entered at the Post Office at Lancaster, Pennsylvania, as second class matter.

Acceptance for mailing at special rate of postage provided for in section 1103, Act of October 3, 1917, authorized September 6, 1926.

Physics-LSA
Johnson
10-14-43
Replaces.

THE
PHYSICAL REVIEW

A Journal of Experimental and Theoretical Physics Established by E. L. Nichols in 1893

VOL. 56, NO. 1 JULY 1, 1939 SECOND SERIES

An Investigation of the Products of the Disintegration of Uranium by Neutrons

PHILIP ABELSON
Radiation Laboratory, University of California, Berkeley California

(Received May 11, 1939)

Products of uranium fission which are chemically antimony, tellurium and iodine have been examined. The following activities have been found:

Antimony	Tellurium	Iodine	Atomic Weight
80 hr.	10 hr.		127
4.2 hr.	70 min.		129
	30 min.	8 days	131
	30 hr.		
5 min.	77 hr.	2.4 hr.	132, 134 or 136
<10 min.	43 min.	54 min.	
<10 min.	60 min.	22 hr.	

Three activities—the 10-hour tellurium, the 70-minute tellurium, and the 8-day iodine—have been identified with known bodies through comparison of half-lives and beta-ray spectra. Genetical relationships of the activities have been established by a technique of periodic separations and by the observation of growth in activity curves.

IN A classical paper in *Die Naturwissenschaften*[1] Hahn and Strassmann announced the discovery of active barium as a product of neutron irradiation of uranium. Meitner and Frisch[2] suggested an explanation of this phenomenon on the basis of the liquid-drop model of nuclei and predicted that the process of uranium fission should be accompanied by the liberation of 200 Mev of energy. Subsequent work in several laboratories[3–7] has demonstrated the production of heavy energetic particles by neutron irradiation of uranium. Further chemical work has shown that the products of uranium fission are numerous. Curie and Savitch[8] have found several rare earth activities and an alkali metal body, and Hahn and Strassman[9, 10] have stated that in addition to a number of barium activities they have found active strontium, yttrium, krypton and cesium. Numerous other workers have also reported activities.

When news of the discovery of the uranium cleavage reached this laboratory, the properties of the x-rays emitted by the 77-hour "trans-

[1] O. Hahn and F. Strassmann, Naturwiss. **27**, 11 (1939).
[2] L. Meitner and O. Frisch, Nature **143**, 239 (1939).
[3] O. Frisch, Nature **143**, 276 (1939).
[4] F. Joliot, Comptes rendus **208**, 341 (1939).
[5] R. D. Fowler and R. W. Dodson, Phys. Rev. **55**, 417 (1939).
[6] R. B. Roberts, R. C. Meyer and L. R. Hafstad, Phys. Rev. **55**, 417 (1939).
[7] G. K. Green and L. W. Alvarez, Phys. Rev. **55**, 417 (1939).
[8] I. Curie and P. Savitch, Comptes rendus **208**, 343 (1939).
[9] O. Hahn and F. Strassmann, Naturwiss. **27**, 89 (1939).
[10] O. Hahn and F. Strassmann, Naturwiss. **27**, 163 (1939).

uranic" activity were immediately examined. The radiations were shown to be characteristic iodine x-rays. Further work established that the 77-hour "transuranic" body is tellurium and that its 2.4-hour daughter is iodine.[11] A series of investigations was then made on those products of uranium cleavage which are chemically antimony, tellurium and iodine. The results[12, 13] have been published. The purpose of this paper is to give in detail a description of the work. The table in the abstract summarizes the activities. The values of half-lives given in this table take precedence over values previously quoted by the author.

Method

Uranium samples were activated by neutrons formed through the bombardment of beryllium by 8.0-Mev deuterons produced by the cyclotron. Samples were placed about one centimeter from the beryllium source and separated from it by a water-cooling brass plate. On all other sides the uranium was surrounded by paraffin. In other words, the uranium was bombarded by a mixture of slow and fast neutrons.

Several different chemical forms of uranium were employed, namely, uranyl nitrate ($UO_2(NO_3)_2 \cdot 6H_2O$), uranyl chloride (UO_2Cl_2) and uranium oxide (UO_3). In most cases the C.P. salts were repurified. Activities were for the most part measured on a Lauritsen type electroscope provided with an aluminum window 0.0005" thick. Calibration with a uranium standard showed one division per second to be equal to one-third of a microcurie. For the identification of x-rays a methyl-bromide-filled ionization chamber was used in conjunction with a d.c. amplifier of the type described by Dubridge. The ionization chamber was also used for the measurement of weak beta-activities.

Chemical Separations

In making chemical separations several factors were considered. First, it was highly desirable to make the separation a specific one. This is not always possible without resorting to a fairly complicated procedure. Second, it was necessary to make the separations in the shortest possible time. Finally, it was desirable to have separations which were semi-quantitative.

The separation employed for antimony was a volatilization of stibine (SbH_3) and the subsequent passage of this gas into a silver nitrate solution. Under these circumstances one obtains a precipitate of silver antimonide (Ag_3Sb).

In the antimony experiments a 10-gram sample of uranium oxide was irradiated with neutrons. After the activation had ceased, the oxide was dissolved in 4 normal sulfuric acid.

The sulfate solution was introduced into a hydrogen generator (employing zinc and 4 normal sulfuric acid). Under these conditions stibine is formed, which is carried along in the stream of hydrogen. The stibine, on coming in contact with a silver nitrate solution, forms a black precipitate of silver antimonide. The only element which reacts similarly is germanium. A later experiment showed that the activity was due to antimony and not to germanium. Besides being an almost specific separation, the stibine volatilization can be performed very quickly. In one experiment it was possible to start reading an antimony precipitate within four minutes after the conclusion of an irradiation. The sole disadvantage of this separation is the fact that only ten percent of the antimony is reduced to stibine. The remainder is precipitated in the hydrogen generator in the metallic state.

In order to separate tellurium from all the other elements, an extensive procedure is necessary. However, a fairly specific precipitation is obtained through the reduction of the element by sulfurous acid. If one passes sulfur dioxide into a boiling 3 normal hydrochloric acid solution containing appropriate carriers, gold, palladium, tellurium and selenium are precipitated completely in the elemental state. A large fraction of the polonium is likewise carried down. At first glance the precipitation of gold, palladium, selenium and polonium seems a great handicap. However, supplementary experiments showed that those four elements were not among the strongly active products of uranium cleavage. Thus, gold and palladium were separated from tellurium by reduction with formic acid. Selenium was precipitated from 12 normal hydrochloric acid by sulfur dioxide. Tellurium is not

[11] P. Abelson, Phys. Rev. **55**, 418 (1939).
[12] P. Abelson, Phys. Rev. **55**, 670 (1939).
[13] P. Abelson, Phys. Rev. **55**, 876 (1939).

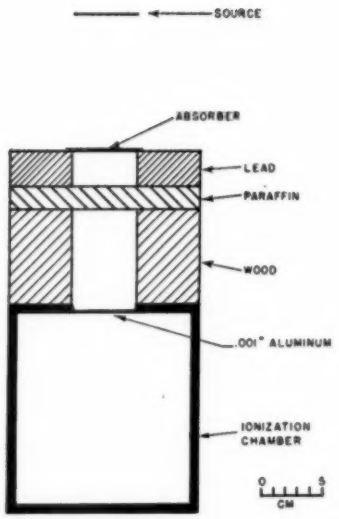


FIG. 1. Experimental arrangement used for determination of x-ray absorption.

precipitated under these conditions. On the other hand, hydrazine hydrochloride precipitates tellurium but not polonium. The reduction of tellurium by sulfur dioxide possesses the further advantages that it is quantitative and can easily be performed in ten minutes.

Several methods were employed for the separation of iodine. In one the iodine was distilled from a dilute nitric acid solution into dilute sodium sulfite. The iodine was then precipitated as silver iodide from a solution acidified by nitric acid.

In a second procedure the iodine was extracted from dilute nitric acid solution with carbon tetrachloride. To insure complete extraction the process was repeated. The iodine was extracted from the carbon tetrachloride by a dilute sodium sulfite solution, and subsequently precipitated as silver iodide. A third procedure—which gave the most satisfactory results in the periodic iodine extractions from tellurium—was carried out in the following manner. The metallic tellurium precipitate was dissolved in nitric acid, to which concentrated sulfuric acid was added, and the mixture was boiled until fumes of sulfuric acid were evolved. The solution was diluted to 6 normal, and silver sulfate was added. The mixture was then saturated with sulfur dioxide. Under these conditions tellurium remains in solution and iodine is reduced from either iodate or iodine to iodide. Any iodine liberated from the tellurium is, therefore, reduced from the valence state in which it was formed to the iodide state in which it is precipitated as silver iodide. The two substances which might interfere in all three of these separations are bromine and element 85. Bromine was eliminated by virtue of the solubility in concentrated ammonia of silver bromide. Element 85 was most easily disposed of by proving that the halogen activities were not genetically related to polonium. The validity of the chemical separations is enhanced by the fact that most of the activities studied possess genetical relationships with other activities. Thus, in the case of the antimony activities decaying into tellurium, if one shows that the daughter substance is uniquely tellurium, one confirms the chemical evidence that the parent is antimony.

For examination of the x-rays emitted by the 3-day tellurium and the 8-day iodine a methyl-bromide-filled chamber was used. The experimental arrangement is shown in Fig. 1. It can readily be seen that the geometrical conditions are excellent. The effect of fluorescent radiation from the absorber is negligible. The cadmium, indium and tin absorbers were used in the form of metallic foils. Antimony as SbO_2 was precipitated in uniform layers on filter paper.

It is well known that the shape of a beta-ray absorption curve is a function of the geometrical conditions employed in obtaining the curve. Hence in showing that one body has the same beta-ray energy as another, it is very desirable to measure the absorption of the radiation from the two substances under the same geometrical conditions. In this work the active samples were of the same thickness and were prepared for examination in similar manners. Furthermore, the intensities of the activities were made nearly equal.

THE 77-HOUR ACTIVITY

Of the fragments of uranium cleavage one of the most thoroughly investigated bodies is a 77-hour activity. Examination of the radiation of this body reveals the emission of x-rays. The emission of these x-rays was first established in March, 1938. Examination of the radiation at that time through use of copper absorbers gave

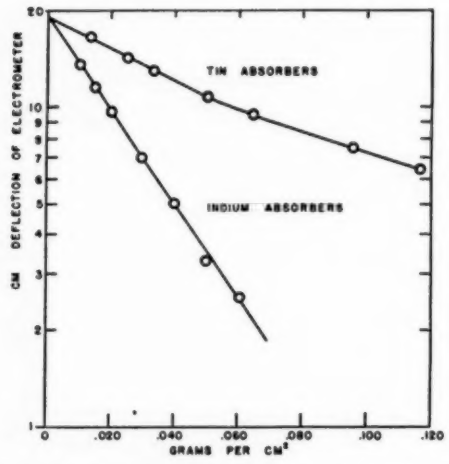


FIG. 2. Absorption of x-rays emitted by 77-hour tellurium. The gamma-background which has been subtracted was about equal to the x-ray effect.

an absorption coefficient which was compatible with the hypothesis that the x-rays were L radiation from a "transuranic" element. The absorption coefficient found was less than expected but geometrical conditions were considered to be poor enough to explain the low result. At that time an electroscope filled with methyl bromide was employed as the detecting instrument, and the available intensity of the 77-hour body was small. When the news of the work of Hahn and Strassmann was heard, the characteristics of the x-rays were at once re-examined. On that occasion a much stronger sample ($\sim$50 microcuries) was available. A number of absorbers were used—copper, bromine, strontium, molybdenum, silver—all of which showed an absorption which increased with atomic number. On trying iodine it was found that the absorption coefficient had markedly decreased. Further experiments revealed an absorption discontinuity lying between tin and indium (Fig. 2). This discontinuity is compatible with the assumption that the x-rays are the characteristic x-rays of iodine. A confirmation of that hypothesis is seen in Fig. 2. Here indium whose K level can be excited by both iodine K_α and K_β shows a single absorption coefficient. On the other hand, tin, whose K level can be excited only by iodine K_β, shows two absorption coefficients in its absorption curve. More recently these x-rays have been investigated by Feather and Bretscher[14] who also find a discontinuity in absorption between indium and tin. They likewise conclude that the radiations are the characteristic x-rays of iodine.

On separating tellurium from a tellurium precipitate which has aged for a week, one can observe a growth in the separated activity (Fig. 3). The activity of the tellurium reaches a maximum about 10 hours after separation. The growth indicates the formation of a daughter substance whose half-life is about $2\frac{1}{2}$ hours. This daughter substance can be easily separated from the tellurium; it has all the chemical properties of iodine, and it decays with a 2.4-hour half-life. The value obtained here for the half-life of the 77-hour body was somewhat different from the 66-hour value quoted by Meitner, Hahn and Strassmann.[15] It was found that a tellurium precipitate made from activated uranium a day after irradiation showed an initial half-life close to 66 hours. In less than a week the slope of the decay curve had become that of a 77-hour activity. In view of the fact that a 30-hour body has been found among the active tellurium fragments, it was assumed that the 66-hour value results from a combination of 30-hour and 77-hour activities. Accordingly, the following experiment was performed. Tellurium was pre-

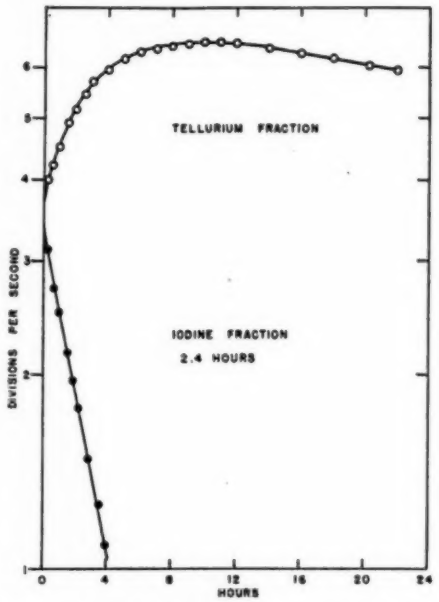


FIG. 3. Decay of 2.4-hour iodine separated from 77-hour tellurium and growth of activity in 77-hour tellurium.

[14] N. Feather and E. Bretscher, Nature **143**, 516 (1939).

[15] L. Meitner, O. Hahn and F. Strassmann, Zeits. f. Physik **106**, 249 (1937).

cipitated from activated uranium at the conclusion of a neutron irradiation. After 12 hours had elapsed, a series of periodic iodide separations was made, and the decay curve of each extract was followed. The iodide activity was found to consist almost entirely of a 2.4-hour body together with a small fraction possessing a half-life of 8 days. The results of these extractions give a decay curve of the parent 77-hour tellurium (Fig. 4).

The 77-hour body with its 2.4-hour daughter constitute a pair of activities which were considered to be "transuranic" elements. The genetical relationships given by Hahn, Meitner and Strassmann[15] list the 77-hour period as the daughter of a 59-minute body. This assignment was investigated thoroughly and found to be in error. The experiment was conducted as follows: A 100-gram sample of uranyl chloride was activated for 30 minutes with neutrons. The salt was dissolved in 3 normal hydrochloric acid, and 100 milligrams of tellurium in the form of tellurous acid were added. To sustain antimony 750 milligrams of antimony chloride were added. Tellurium was precipitated at once by passing sulfur dioxide into the boiling solution. A series of precipitations of tellurium was made at 10-minute intervals with 100 milligrams of the carrier being added before each precipitation. Decay curves of the active tellurium precipitates were followed and the relative amounts of the 77-hour activity were found both by analysis of the beta-decay curves and by a measurement of the x-rays. The results are given in Table I. It is evident that the parent of the 77-hour tellurium must have a half-life of much less than the 59-minute value given by Meitner, Hahn and Strassmann. In fact, the apparent half-life of the longest-lived ancestor is 5 or 6 minutes. Moreover, any lack of completeness of the precipitation of the tellurium would lengthen the apparent half-life of the parent rather than shorten it. Weighing of the tellurium precipitate showed that the precipitations were practically complete.

TABLE I. *Decay of active tellurium precipitates.*

	TIME AFTER START OF IRRADIATION	ACTIVITY OF 77-HOUR COMPONENT
Conclusion of irradiation	30 min.	
1st tellurium precipitation	40 min.	3.4 div./sec.
2nd tellurium precipitation	50 min.	0.25 div./sec.
3rd tellurium precipitation	60 min.	0.07 div./sec.
4th tellurium precipitation	70 min.	0.02 div./sec.

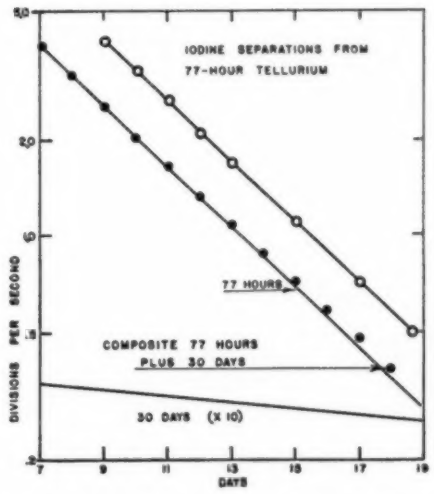


FIG. 4. Determination of half-life of 77-hour tellurium.

It is established, then, that the 77-hour activity is formed from a body whose quantity decays with a half-life of about five minutes. It must be borne in mind that this does not mean that the parent antimony is necessarily a 5-minute body. On the other hand, for instance, a 5-minute tin might decay into a shorter-lived antimony. On making the successive tellurium separations, one would find an apparent half-life of 5 minutes for the parent of the 77-hour activity. However, a 5-minute antimony activity has been found which probably is the parent of the 77-hour substance.

It is quite possible that the 2.4-hour body is the final active member of a chain of beta-decays. There are three stable xenon isotopes—132, 134 and 136—which could terminate the series of activities. If the 2.4-hour body has a daughter substance, the half-life of that activity must be greater than several months or shorter than one minute. The evidence for the above statement follows: If one makes a speedy iodine separation from the 77-hour activity, one finds that the 2.4-hour activity decays following a simple exponential law into the background. A sample of iodine whose intensity was twenty divisions per second was placed on the electroscope three

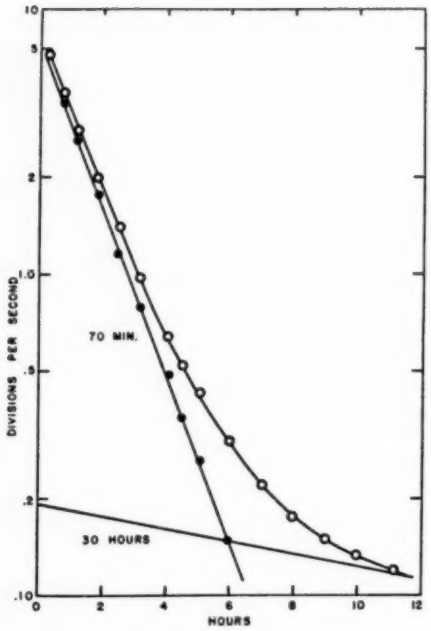


FIG. 5. Decay curve of tellurium separated from parent antimony activities showing the 70-minute daughter of a 4.2-hour activity.

minutes after the beginning of an iodine separation. If the half-life of the daughter were in the range from one minute to 100 days, the curve would not have been a simple exponential.

THE 70-MINUTE ACTIVITY

Among the active fragments of uranium fission is an antimony isotope whose half-life is 4.2 hours. If an activated uranium sample is allowed to age for six hours after termination of bombardment and then antimony is separated, the activity curve shows a growth. The intensity reaches a maximum in two hours. This indicates the growth of a tellurium daughter whose half-life is of the order of an hour. Separation of tellurium from the antimony revealed that the tellurium had a half-life of 70 minutes (Fig. 5). The determination of the parent half-life in the antimony is complicated by the presence of other long-lived antimony isotopes which themselves produce tellurium isotopes. Hence, the most accurate determination of the half-life of the 4.2-hour body is obtained through making periodic extractions of tellurium from antimony and then following decay curves of the separated tellurium fractions to determine the amount of seventy-minute body present at the time of extraction. The results are shown in Fig. 6. The equality of the value of the 70-minute half-life to that found by Seaborg, Livingood and Kennedy[16] for the lower isomer of Te^{129} suggested that the 70-minute body produced from uranium might be Te^{129}. Absorption curves of the beta-rays emitted by both of the tellurium isotopes have been obtained. The known Te^{129} was prepared from a sample of tellurium which had

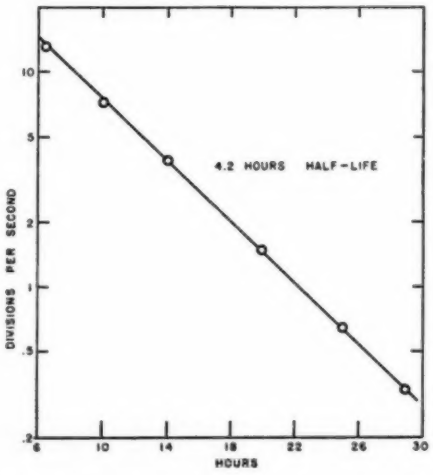


FIG. 6. Decay of a 4.2-hour antimony activity established through periodic tellurium separations of the daughter substance.

been bombarded by deuterons. A mixture of the 70-minute Te^{129} and the 10-hour Te^{127} was obtained separately from the parent isomers of 30-day and 90-day half-lives by means of an extraction devised by Seaborg, Livingood and Kennedy.[16] A beta-ray absorption curve was taken immediately after separation. The decay curve of the mixture of 70-minute and 10-hour bodies was followed and the fraction of each present thus determined. Fifteen hours after the isomer separation with the 70-minute body completely gone a second beta-ray absorption curve was followed. The characteristics of the beta-rays of the 10-hour body were thus obtained. The true absorption characteristics of the 70-minute body were then found by correcting the first absorption curve for the effect of the presence of the 10-hour activity. Actually this correction was very small because the 10-hour

[16] G. T. Seaborg, J. J. Livingood and J. H. Kennedy, Phys. Rev. **55**, 794 (1939).

body had an activity only five percent of that of the 70-minute body. The absorption curves of the 70-minute bodies obtained from both uranium and tellurium are shown in Fig. 7. It is seen that the two curves are parallel over a factor of 100 in intensity. It seems probable then that the two 70-minute bodies are identical.

The 10-Hour Activity

Another cleavage product of uranium which can be identified with a known activity is the 10-hour tellurium. Beta-ray absorption curves of this body and of the lower Te^{127} isomer[16] are identical. These curves are shown in Fig. 7. The 10-hour substance is obtained as the daughter of an 80-hour antimony isotope. If antimony is

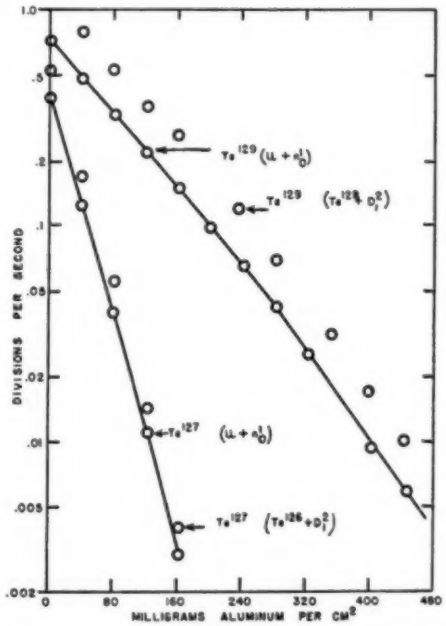


Fig. 7. Beta-ray absorption curves of 70-minute Te^{129} and 10-hour Te^{127}.

precipitated from an active sample which has been allowed to age for a week, the decay curve shows a growth which indicates the production of a daughter substance whose half-life is of the order of 10 hours (Fig. 8). On separating tellurium from the active antimony a 10-hour tellurium body was obtained. Furthermore, a series of tellurium precipitations showed that the 10-hour daughter was growing from an 80-hour parent (Fig. 8). No experiments have yet been made to determine the parent of the 80-hour antimony.

The 8-Day Activity

Still another fragment of uranium disintegration which may be identified with a known activity is an 8-day iodine body. The decay curve of this substance is shown in Fig. 9. An 8-day iodine activity can be produced through deuteron or neutron activation of tellurium and has been shown to be I^{131}.[16] Several lines of evidence indicate that the two 8-day iodine bodies are identical. Beta-ray absorption curves have been made on both bodies. These are identical. Further evidence is the fact that both of the 8-day iodines were found to emit characteristic xenon x-rays. It has previously been shown that I^{131} emits a gamma-ray of 0.4-Mev energy. Corresponding to any given geometrical arrangement of a source of I^{131} with respect to the ionization chamber, a certain value of the ratio of the gamma-ray to x-ray ionization was observed. When a source of the 8-day body obtained from uranium was placed in the same geometrical configuration as that of the I^{131}

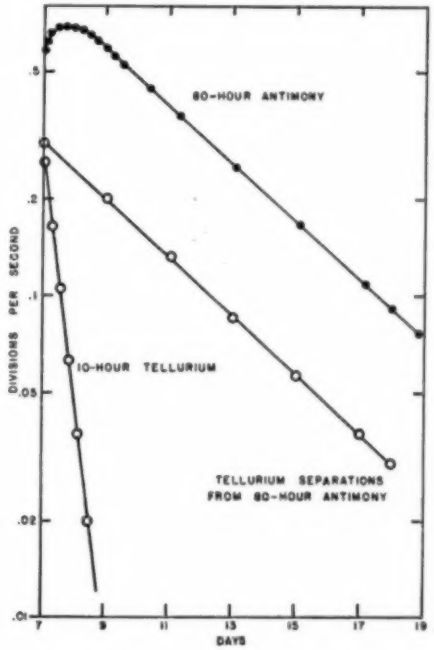


Fig. 8. Decay of 10-hour tellurium daughter extracted from 80-hour antimony and growth of activity in the 80-hour mother substance. The half-life of the antimony activity was also determined by periodic tellurium extractions of its daughter substance.

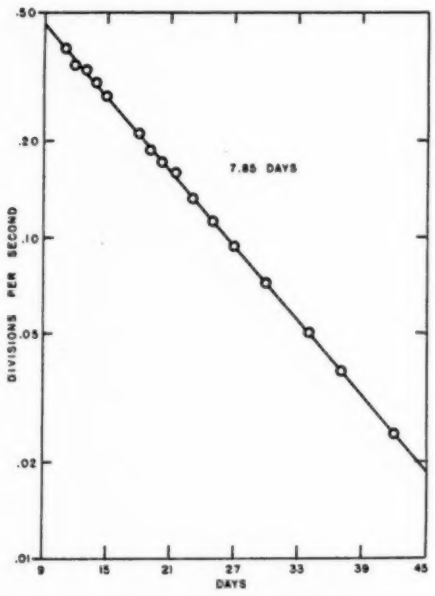


FIG. 9. Decay curve of 8-day iodine.

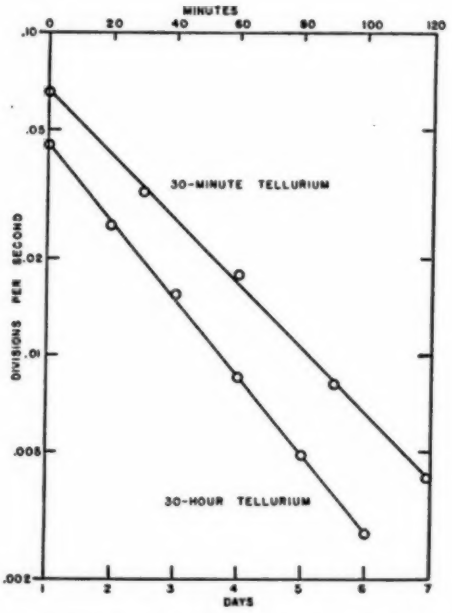


FIG. 10. Determination of half-lives of tellurium parents of 8-day iodine through periodic separations of iodine from tellurium.

identical values of the ratio of gamma-ray to x-ray intensities were found.

It has been shown[16] that Te^{131} possesses two isomers of 30-hour and 25-minute periods with the higher energy state belonging to the 30-hour body. Activities of these same half-lives have been found in tellurium precipitated from activated uranium. The periods of these parent tellurium activities were established by making a series of periodic iodine extractions from the precipitated tellurium. The results are shown in Fig. 10. In both cases the first iodine extract was discarded because some iodine could be carried down with the original tellurium precipitate. These experiments show that a 30-hour and a 30-minute Te^{131} are formed by neutron irradiation of uranium. Presumably these bodies are identical with those produced by activation of tellurium itself.

Two other tellurium periods were demonstrated through the technique of repeated iodine separations from tellurium which had been precipitated from neutron activated uranium. The decay curve of a typical iodine extract is shown in Figs. 11 and 12. This curve can be analyzed into two main components of 54-minute and 22-hour half-lives. A series of such iodine separations was made at 30-minute intervals and the activity of each extract was followed. The

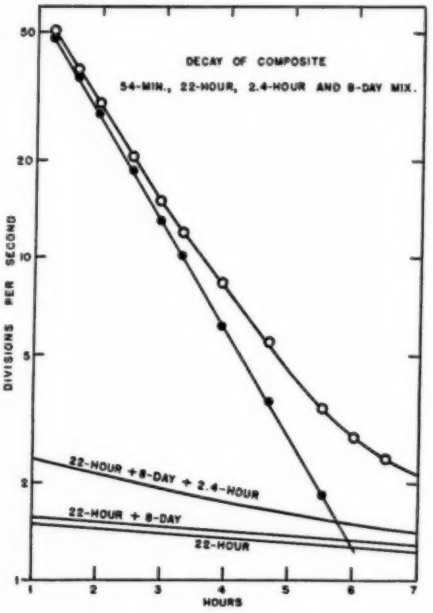


FIG. 11. Decay of a typical mixture of iodine activities separated from tellurium. The solid dots correspond to the 54-minute activity.

presence of the 2.4-hour and the 8-day bodies was taken into account. This is especially important in the case of separations made several hours or longer after bombardment. From these extractions it was determined that the 54-

minute iodine grows from a 43-minute tellurium and that the 22-hour iodine is the daughter of a 60-minute tellurium. Results are shown in Fig. 13. An experiment was made to determine the half-lives of the parents of the two tellurium activities. A series of tellurium precipitations was made at 10-minute intervals from an activated uranium sample. After an hour had elapsed these precipitates were dissolved and iodine extracted from them. The first two tellurium separations were discarded. Examination of the iodine activity removed from the third and subsequent precipitates revealed that neither the 22-hour nor the 54-minute body was present. This indicates that the grandparents (if any) of these two iodine substances possess a half-life less than 10 minutes.

A glance at the tellurium activities which have been found reveals four whose half-lives are so close together as to be unidentifiable when lumped together in a single decay curve. The technique of periodic separations of daughter substances has proved successful in solving this problem. However, in order to follow decay curves of the extracts it is necessary to have large initial intensities. The writer feels that the accomplishment of this work is due in large measure to the huge neutron intensities of the cyclotron which have been made available to him.

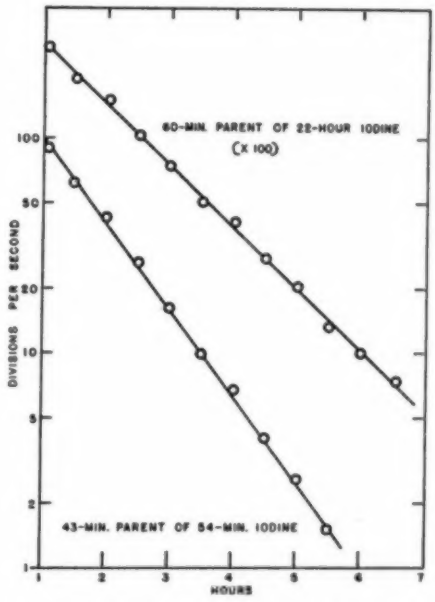


FIG. 13. Determination of half-lives of the 43-minute and 60-minute tellurium bodies through periodic separations of their daughter iodine activities.

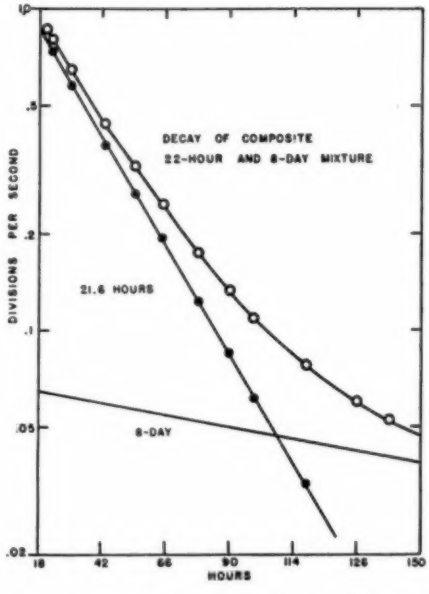


FIG. 12. Decay of a typical mixture of iodine activities separated from tellurium. The 54-minute and the 2.4-hour activities have practically disappeared.

ACKNOWLEDGMENT

Credit is due to Dr. E. Segrè and Dr. G. T. Seaborg for their advice and suggestions, and to Dr. J. G. Hamilton for his assistance in making x-ray and beta-ray absorption measurements.

Thanks are especially due to Professor E. O. Lawrence for his interest in this work and to the Research Corporation for financial assistance.

The Intensity of Neutrons of Thermal Energy in the Atmosphere at Sea Level

C. G. MONTGOMERY AND D. D. MONTGOMERY
Bartol Research Foundation of the Franklin Institute, Swarthmore, Pennsylvania

(Received May 15, 1939)

The number of alpha-particles released in the disintegration of boron by neutrons in the atmosphere was measured in an ionization chamber filled with boron trifluoride. Assuming a cross section for the $n-\alpha$ reaction of 3×10^{-21} cm², the flux of neutrons of thermal energy was found to be 0.091 ± 0.007 per square centimeter per minute, or one thermal neutron for every 16 ionizing cosmic rays. The consequences of the formation and absorption of this intensity of neutrons are discussed.

A NUMBER of investigators[1] have reported the presence of neutrons in the atmosphere, both at sea level and at higher elevations, but no reliable reduction of the observations by applying a correction for the efficiency of detection has been made to determine the actual number of neutrons present. It has been supposed that the neutrons originate in some interaction of the cosmic radiation with matter, the evidence for this consisting of the association of neutrons with cosmic-ray showers and the rapid increase in the number of neutrons at higher elevations. The experiments described below were made in an effort to establish the number of neutrons of thermal energy present in the atmosphere at sea level.

Neutrons were detected by utilizing the alpha-particles released in the disintegration of the boron isotope of mass ten. An ionization chamber in the form of a vertical cylinder of copper 40 cm long and 7 cm in diameter was enclosed in a glass envelope and filled with boron trifluoride gas at atmospheric pressure. The collecting electrode was a fine wire through the center of the cylinder. A dry battery supplied a steady potential of 450 volts to collect the ions formed. The collecting electrode was directly connected to the grid of an FP54 Pliotron. The voltage pulse from the FP54 resulting from a spurt of ionization caused by an alpha-particle was then applied to a type 38 vacuum tube for additional amplification, and the output of this tube was capacitatively coupled to a galvanometer. The position of the galvanometer spot was continuously recorded on moving photographic paper, and the alpha-particles appeared as kicks whose heights were proportional to the amounts of ionization produced by the alpha-particles. The natural alpha-particle background of the chamber caused by radioactive impurities was determined by shielding the chamber with a thick layer of borax which absorbed all neutrons of thermal energy falling upon it. The size-frequency distribution of the alpha-particle pulses so determined is shown in curve *A*, Fig. 1. When the borax shield was removed a larger number of alpha-particles was observed, the increase being the number of alpha-particles resulting from the disintegration

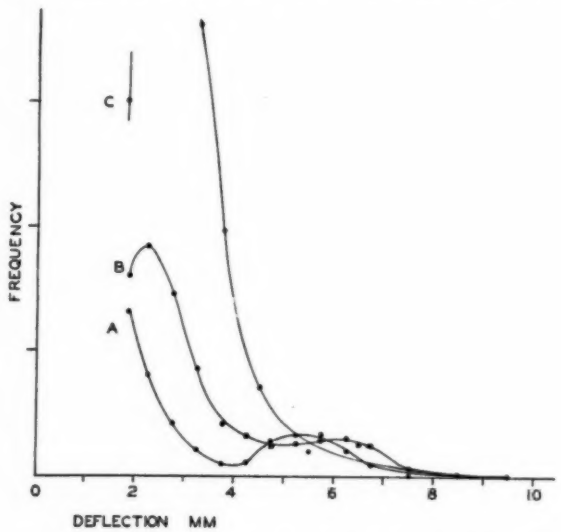


FIG. 1. Size-frequency distributions of the alpha-particles in the ionization chamber. Curve *A*—thick borax shield; curve *B*—no shield; curve *C*—200-milligram Ra-Be neutron source 15 meters from apparatus.

[1] G. L. Locher, Phys. Rev. **44**, 779 (1933); J. Frank. Inst. **216**, 673 (1933); *ibid.* **224**, 555 (1937); L. H. Rumbaugh and G. L. Locher, Phys. Rev. **49**, 855 (1936); E. Fünfer, Zeits. f. Physik **111**, 351 (1938); E. Schopper, Naturwiss. **25**, 557 (1937); D. K. Froman and J. C. Stearns, Phys. Rev. **54**, 969 (1938); H. von Halban, L. Kowarski, and M. Magat, Comptes rendus **208**, 572 (1939).

of the boron nuclei. The size-frequency distribution without the shield is shown in curve *B*, Fig. 1. The distribution of the disintegration alpha-particles, obtained by taking the difference between the two observed distributions, is shown in Fig. 2. The peak is quite broad, and gives definite evidence for the occurrence of several groups of alpha-particles although there is not sufficient resolution to separate the individual peaks. This is in agreement with the recently reported work of W. Maurer and J. B. Fisk.[2]

As an additional check, some observations were made with a radium-beryllium source of neutrons. A 200-milligram source was placed in another room of the laboratory at a distance of about 15 meters, and shielded with about four inches of lead. The alpha-particle distribution in the ionization chamber in the presence of this source is given as curve *C* in Fig. 1. The fact that both curve *B* and curve *C* differ from curve *A* only for the smallest sizes of the ionization spurts, and in an exactly similar manner, can be taken as good evidence that the disintegration of boron by neutrons is responsible for the observed effects. Table I shows the rates of occurrence of the alpha-particles in the different experiments.

In order to estimate the efficiency for detection of neutrons of thermal energy, it is necessary to know the cross section for the reaction involved. It has generally been supposed that the large absorption of boron for slow neutrons was the result of the reaction producing the alpha-particles, and that the contribution of other processes to the absorption was negligible. In that case it would be possible to determine the cross section from the absorption experiments. However, Maurer and Fisk[2] have recently indicated that the $B^{10}(n-p)Be^{10}$ reaction may be

TABLE I. *Rate of occurrence of alpha-particles in ionization chamber.*

ARRANGEMENT OF CHAMBER	TOTAL NUMBER OF ALPHA-PARTICLES PER HOUR	NUMBER OF DISINTEGRATIONS PER HOUR
Borax shield	99± 4	
No shield	190± 6	91± 7
No shield with neutron source present	504±37	405±37
Paraffin shield	177± 8	78± 9

[2] J. B. Fisk, Phys. Rev. **55**, 1117(A) (1939).

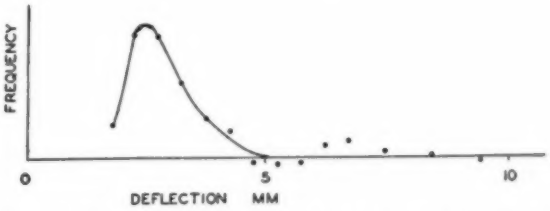


FIG. 2. Size-frequency distribution of the alpha-particles from the disintegration of the boron by neutrons in the atmosphere. This curve is the difference between curves *B* and *A* of Fig. 1.

quite as important an absorption process as the $B^{10}(n-\alpha)Li^7$ reaction. If this proves to be the case, then further experiments must be performed to determine the cross section. Pending these further investigations, we shall therefore adopt the cross section given by Livingston and Bethe[3] determined from absorption experiments which resulted in a value of $\sigma=3\times10^{-21}$ cm^2. The results of the experiments reported here may be easily corrected for a change in this cross section, if necessary. To calculate accurately the efficiency of the ionization chamber it would be necessary to know the angular distribution of neutron velocities. Fortunately, however, the absorption by the gas in the chamber is small, and the neutron flux may be expressed, with an error of less than five percent, as $n/N\sigma V$, where N is the number of B^{10} atoms per cubic centimeter, V the volume of the chamber, and n the number of disintegrations observed. It is necessary to correct for the absorption (about 25 percent) by the boron in the glass envelope of the chamber. Using this expression, we find that the neutron flux, that is, the number of neutrons passing through a sphere of one square centimeter cross-sectional area in unit time is 0.091 ± 0.007 cm^{-2} min.$^{-1}$. The flux of ionizing cosmic-ray particles is[4] 1.48 cm^{-2} min.$^{-1}$, so that we have one neutron of thermal energy for every 16 cosmic rays. The only other observations which allow an estimate of the actual number of neutrons present are those of Fünfer.[1] It is possible to calculate the efficiency of Fünfer's proportional counter, from the data given, to within a factor of about two. A correction of his

[3] M. S. Livingston and H. A. Bethe, Rev. Mod. Phys. **9**, 340 (1937).

[4] J. C. Street and R. H. Woodward, Phys. Rev. **46**, 1029 (1934).

results in this manner gives a neutron intensity in agreement within this limit with the present experiments.

In order to detect the presence of neutrons with more than thermal energy, observations were made with a six-centimeter paraffin shield around the ionization chamber. The size-frequency distribution of alpha-particles observed was similar to curve *B*, Fig. 1, and the number of particles is given in Table I. It is seen to be only a little less than the number observed with no shield. A shield of the thickness used would decrease the number of thermal neutrons to less than ten percent. We must, therefore, conclude that thermal neutrons are produced in the paraffin shield either directly by the cosmic radiation or by the slowing down of fast neutrons.

It is of interest to inquire into the possible mode of formation and the disappearance of neutrons in the atmosphere. Although the experiments described above were performed inside the laboratory, for lack of more complete information, we may take them to be representative of the conditions in the free atmosphere. If this is the case, then the most probable process for the absorption of thermal neutrons will be the $N^{14}(n-p)C^{14}$ reaction. The cross section for this reaction[5] is 11.3×10^{-24} cm^2, which corresponds to an absorption coefficient of 4.8×10^{-4} cm^{-1} of standard air. Now since equilibrium exists, as many neutrons must be formed as disappear, and we must have 4.4×10^{-5} neutron of thermal energy produced per cubic centimeter per minute. If no other process is important for neutron absorption, then this figure must also represent the total production of neutrons of all energies. If we assume that all atoms in the air interact with all the ionizing cosmic rays, we may compute the cross section for the production of a neutron to be 5.5×10^{-25} cm^2. This cross section is surprisingly high, and indeed, since other absorption processes may be important, and since all cosmic rays probably do not produce neutrons, this cross section must represent only a lower limit to that of the true process. It seems likely that the neutrons are produced by high energy photons, since they seem to be associated at sea level with showers,[6] and increase with elevation as do the large bursts of ionization.[7] It is interesting to note in this connection that the nuclear disintegrations observed by Anderson and Neddermeyer[8] in cloud chambers increase with elevation in the same way, and quite probably give neutrons as disintegration products as well as the heavily ionizing particles which are observed. The total energy loss of the cosmic radiation in producing neutrons is probably negligibly small, since if we take 10^7 electron volts as the energy necessary to free a neutron from a nucleus, the energy loss in air by each cosmic ray would only be 2.3×10^5 electron volts per meter of water equivalent. The number of protons resulting from the absorption of neutrons is likewise small. Since the protons will have a range of only about one centimeter of air, there will be only one proton for every 31,000 ionizing cosmic rays crossing a given area. Such a frequency of protons would be observable in a cloud chamber if it were possible to distinguish them from the radioactive contamination.

Although the experiments described here do not represent the most precise determination that is possible, we believe that they give the number of neutrons of thermal energy at sea level to within about ten percent, an accuracy sufficient to permit an evaluation of the importance of neutrons in cosmic-ray phenomena. Further experiments are under way.

[5] J. R. Dunning, G. B. Pegram, G. A. Fink and D. P. Mitchell, Phys. Rev. **48**, 265 (1935).

[6] G. L. Locher, reference 1.

[7] E. Fünfer, reference 1.

[8] C. D. Anderson and S. H. Neddermeyer, Phys. Rev. **50**, 263 (1936).

The Self-Diffusion of Copper*

J. STEIGMAN,†
Columbia University, New York, New York

W. SHOCKLEY AND F. C. NIX,
Bell Telephone Laboratories, New York, New York

(Received May 6, 1939)

The self-diffusion of copper has been measured with a radioactive isotope prepared by neutron bombardment of zinc. The isotope is electroplated in a thin layer upon the surface of a copper disk. By measuring the decrease in β-ray emission from the sample due to the penetration of the radioactive copper, the self-diffusion constant is determined within the limits of accuracy as 11 cm²/sec. $\exp(-57{,}200/RT)$. Methods of analyzing such data are given and the significance of the variation of activation energy between self-diffusion and heterogeneous diffusion is discussed.

PART I

DIFFUSION is a process of fundamental significance in practical and theoretical metallurgy.[1] There are, however, no quantitative theories for diffusion and there is even some question as to the exact qualitative nature of the diffusion in metals. Three mechanisms are regarded as being possible: atomic interchange occurs between neighboring atoms; some atoms occupy interstitial positions and diffuse by moving from one such position to another; "holes" or vacant lattice sites exist and diffusion results from atoms moving from site to site through the medium of the holes.

The simplest type of diffusion is self-diffusion. In the strict sense of the term self-diffusion implies diffusion of something into its exact counterpart; however, used more loosely it implies diffusion of one isotope into another isotope of the same chemical element. The earliest experiments in self-diffusion in metals were performed by von Hevesy and his collaborators[2–4] using ordinary lead into which was diffused thorium B which is a radioactive isotope of lead. By analyzing the distribution of the radioactive isotope in the lead, they were able to determine the rate of self-diffusion. Now that artificial radioactive isotopes are available it is possible to investigate many additional cases of self-diffusion.

The present paper is concerned with the self-diffusion of copper. An indirect method of determining the rate of diffusion has been used, the procedure being as follows: upon a copper disk is electroplated a very thin layer of radioactive copper and the radioactivity is measured with a Geiger-Müller counter. The disk is then heated in vacuum in such a manner as to cause diffusion to take place without appreciable loss from evaporation. The penetration of the radioactive atoms below the surface of the disk results in an absorption in the disk of their radioactive radiation and thereby produces a decrease in radioactivity as observed by the Geiger-Müller counter.

After the stopping power of copper for the radioactive radiation is determined, it is possible to use the observed activities in order to determine the depth of diffusion of the active isotope and hence to measure the rate of diffusion. Part II of this paper gives the methods of producing and concentrating the radioactive isotope and of measuring the activities and stopping power of copper. Part III describes the technique used in the diffusion process and Part IV presents the theoretical analysis and the results.

* The results given in this paper were presented at the 40th annual meeting of the American Physical Society, Washington, D. C., December 27–29, 1938. Publication assisted by the Ernest Kempton Adams Fund for Physical Research of Columbia University.

† Now in the Chemistry Department at the City College of the College of the City of New York.

[1] R. F. Mehl, J. App. Phys. **8**, 174 (1937).

[2] J. Groh and G. v. Hevesy, Ann. d. Physik **63**, 85 (1920).

[3] G. v. Hevesy and A. Obrutscheva, Nature **115**, 674 (1925).

[4] G. v. Hevesy and W. Seith, Zeits. f. Physik **56**, 790 (1929).

Part II

Radioactive copper (Cu^{64}—period 12.8 hours) was prepared from zinc (Zn^{64}) by neutron bombardment with a radon-beryllium source of neutrons varying in strength from 300 to 600 mc. Ninety grams of zinc (C. P. Eimer & Amend product) was irradiated for about 13 hours and then treated with concentrated hydrochloric acid so as to dissolve all but several hundred milligrams, which were found to retain practically all the activity. This small residue was then dissolved in hydrochloric acid; 30 cc of concentrated sulphuric acid was added; and the mixture evaporated until fumes of SO_3 were given off. The resultant solution was next diluted to 400 cc and filtered. The radioactive copper was electroplated from the filtrate upon a rapidly rotating disk cathode of copper, of 3 cm diameter and 0.5 mm thickness.[5] The operating current was 11–12 milliamperes, the voltage 2.0–2.2 volts, the speed of rotation 4500 r.p.m. and the period of electrolysis 30–45 minutes.

The copper disks which were used were rolled from bars of spectroscopically pure copper which was obtained from Adam Hilger & Company, London (referred to as "Hilger" copper). In addition, so-called "Chile" copper of lesser purity was also used. It was found that the very small traces of radioactive copper could not be successfully plated on "Hilger" copper no matter how clean the surface. Even when the latter had received a careful metallographical polish, very little radio-copper would deposit. Finally, it was found that buffing the surface of such a disk with rouge on a rapidly rotating cloth wheel produced a surface upon which radioactive copper would deposit satisfactorily. The plating was facilitated by adding to the plating bath a drop of "reagent" copper sulphate solution (about 0.1*N*) to act as a carrier. The radioactive copper was more readily plated upon the "Chile" copper and ordinary sheet copper than upon Hilger copper, presumably because of small amounts of impurities in them. It is possible that the buffing operation added minute amounts of impurity or produced strains in the metal which would have the virtual effect of metallic impurities.

Spectrochemical analyses were made of a large number of the disks. In some cases traces of Zn (less than 0.01 percent) were found and in other cases no detectable amount was present. There was no apparent correlation between the finding of zinc and the degree of diffusion. Traces of Sn, Pb, Mg, Si, Al and Fe were found in all analyses.

After the disk was plated with radio-copper, it was cut into four equal segments, each one marked in a characteristic manner. The activities of the segments were then measured, either individually or collectively, in a special holder which fitted around a thin silver-walled Geiger-Müller tube. This holder was constructed so as to have positions for holding four quadrants evenly spaced in definite positions about the counting tube. It was found that successive measurements of the activity of a copper segment placed in the same position in the holder several times did not differ from each other by more than one percent. This arrangement permitted (1) accurate measurement of any single quadrant, and (2) the simultaneous measurement of four quadrants—a feature of value at the end of a long run when the activity of any one sample was too near to the background value for any precise determination.

The radioactivity was measured by means of a thin-walled glass counter, whose cathode was of chemically deposited silver (now made by Eck & Krebs, New York). Such counters have a high sensitivity (ratio of count surrounded with KCl to background count about 40 to 1), and a record of constant performance over a period of several years. The counter was connected to a modified Neher-Harper circuit,[6] and a thyratron de-multiplying recorder system.

The background of this counter was about 17 counts per minute. In general, a sufficient number of counts per minute per sample were taken so that the precision was of the order of magnitude of three percent to 1.5 percent (i.e., 1000 to 4000 counts in all were taken for any one reading). The initial activities ranged from 800 counts per minute to 100 counts per minute. The final activities ranged from 15 counts per minute above background to several hundred counts per minute per quadrant above back-

[5] J. Steigman, Phys. Rev. **53**, 771 (1938).

[6] H. V. Neher and W. W. Harper, Phys. Rev. **49**, 940 (1936).

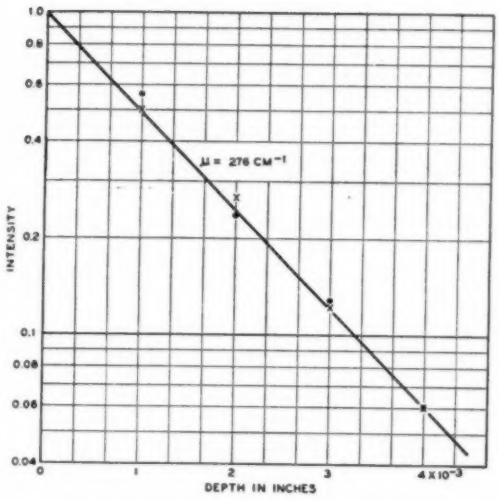


FIG. 1. Variation of transmitted activity as a function of the thickness of the interposed copper.

ground. In order to eliminate the effects of possible fluctuations in the counting rate, a radioactive disk, prepared at the same time as the disks used for the diffusion, was set aside as a control standard. During any particular diffusion run, the activity of this disk was measured immediately before or after the sample was measured, and in only one case did the activity of the control disk vary markedly beyond the estimated precision from the activity calculated on the basis of a 12.8-hour period.

The effective absorption coefficient of copper for the beta-rays from radio-copper was determined by placing a disk upon whose surface radio-copper had been plated in a horizontal position below the Geiger-counter, and measuring the transmitted intensity with various thickness of copper foil interposed between the disk and the counter. Plots of the log of the activity against thickness of the foil yielded a straight line. A line with the same slope was obtained for a thin activated foil of copper which was wrapped around the counter tube with various thicknesses of inactive foil interposed. The results are shown in Fig. 1. The transmission is decreased approximately by a factor of $\frac{1}{2}$ for each mil of copper interposed between the disk and the counter and the corresponding value of the absorption coefficient μ is 276 cm^{-1}.

The effect of diffusion is determined by measuring the activity of the material before diffusion and after diffusion. In interpreting the results it is necessary to know what part of the decrease in activity is due to diffusion and what part to the ordinary decay. Since the period of decay is known to be 12.8 hours the effect of time decay can be eliminated by multiplying the second measurement by exp $(t/12.8)$ where t is the time in hours between measurements. The remaining decrease is due to diffusion and is expressed by stating the fraction, F, of the original activity which is observed after diffusion. The values of F found for runs of various times and temperatures are shown in Figs. 4, 5 and 6.

PART III

In order to obtain measurable amounts of diffusion, it was necessary to maintain the activated quadrants of the disks at elevated temperatures for known times. The quadrants were placed in a special copper holder constructed in a manner permitting the active or plated surfaces of the quadrants to be firmly held between the two halves of the holder with the radioactive surface of one quadrant opposite that of another. This relatively large copper holder also served to maintain the several quadrants at the same temperature. The placing of the quadrants in the "face-to-face" position tended to reduce the net loss of radioactive copper atoms in that atoms evaporating from one quadrant would tend to deposit on the opposite one. As is shown in the appendix, the transfer of radioactive material from the face of one quadrant to the other has no effect on the value of F provided that one deals with the sum of activities of the two quadrants rather than with each individually. However, any net loss from the two quadrants would be a source of error and for this reason checks were run at various times to determine the amount of activity lost by evaporation to the holder and other parts of the apparatus. The activity found in the other parts was always within the experimental error.

The holder with copper disks was then placed in a fused quartz tube and evacuated to a pressure of less than 10^{-4} mm of Hg. In some instances the quartz tube was sealed off while in others the tube was pumped continuously throughout the diffusion experiment. We were unable to detect any

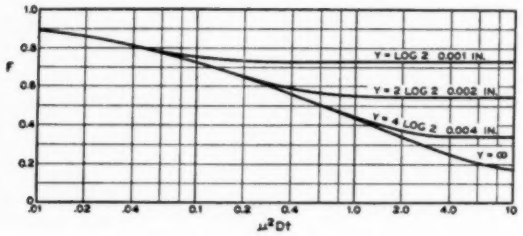


FIG. 2. Apparent activity from plated face *versus* $z=\mu^2Dt$.

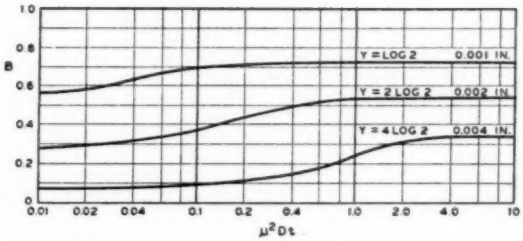


FIG. 3. Apparent activity from back face *versus* $z=\mu^2Dt$.

difference in the diffusion rates between the runs carried out in the two kinds of tubes, i.e., between tubes sealed off after evacuation and continuously pumped tubes. The temperature was measured by means of platinum-platinum ten percent rhodium thermocouples. In the case of the continuously pumped tubes the protected hot-junction was placed in a hole in the copper holder. This permitted us to obtain accurate measurements of the temperature in the immediate vicinity of the diffusion specimens. In the case of the sealed-off quartz tube the thermocouple junction was placed on the outside of the tube near the place occupied by the copper holder. In the latter method, with the sealed off tubes, checks showed the temperature of the specimen to be accurate to within ± 2°C. Corrections were made to offset the lag in heating of the holder with specimens in the sealed off tubes. The temperature of the furnace was maintained constant either by means of a voltage regulator or a potentiometric temperature controller. The temperature was held constant to within ± 2°C. On completion of a run the quartz tube was immersed in a bath of cold water in order to terminate the diffusion run at the desired time. Examination of the copper disks after the diffusion runs showed them to be clean and completely free of any oxide film. Runs were made at three temperatures, 750°, 850° and 950°C and the results are shown in Figs. 4, 5 and 6. In some cases the points shown represent the average for all four quadrants. In other cases, two points are shown connected by a line; in these, each point corresponds to a pair of quadrants facing each other. It is seen that the spread between different runs is in general larger than the spread in an individual run. At present we have no satisfactory explanation for the large spread. Tests have been made using two quadrants from one disk in one run and the remaining two quadrants in a different run; also tests have been made using both the Hilger and Chile copper; these show that there is no consistency between the type of disk and the variations of the final results. There is no apparent correlation with the grain size obtained after diffusion as determined by microscopic examination. Some preliminary experiments with tubes containing hydrogen gas suggest that there may be an important correlation between the atmosphere and the diffusion and we hope to experiment further on this effect in the future.

PART IV

According to the considerations presented in the last section, which are justified in the Appendix, the effects of evaporation from one quadrant to the other can be disregarded if the total activity of the two quadrants is dealt with. For this reason we may consider a single quadrant. Let us suppose that the quadrant is of thickness w, that the absorption coefficient is μ, and that the radioactive material which was initially on the surface of the quadrant has diffused for a time t, the diffusion constant being D. By the use of standard mathematical methods one can find, as is shown in the Appendix, how the fraction of the activity F depends upon these quantities. One finds that F is a function of two variables, z and y where

$$z=\mu^2Dt \tag{1}$$

and

$$y=\mu w. \tag{2}$$

In Fig. 2 we give the results of these calculations, each curve corresponding to a fixed thickness or y value.

It is possible to carry out an exactly similar procedure for the case where the measurement of the activity is made from the back of the

quadrant. In this case we denote by B the fraction of the initial activity as measured from the plated side which is later observed from the back side. In Fig. 3 are given a series of curves like those in Fig. 2.

For the conditions of our experiments the disks are thick in the sense that very little radiation can get through them and furthermore, the depth of diffusion is small compared to the thickness—mathematically $y>10$ and $y^2>3z$. Under these conditions F is well approximated by

$$F(z, y)=F(z, \infty)=e^z[1-P(\sqrt{z})], \qquad (3)$$

where

$$P(\sqrt{z})=\frac{2}{\sqrt{\pi}}\int_0^{\sqrt{z}} e^{-v^2}dy \qquad (4)$$

is the error integral. From a measurement of F we can use our curve to determine z, and then using the known time of diffusion and the value for μ we can calculate D from Eq. (1). A more satisfactory procedure, and the one used, is as follows: at a given temperature D is a constant, hence a plot of F *vs.* z on a logarithmic scale should be the same as a plot of F *vs.* t displaced by log $(\mu^2 D)$. Figs. 4, 5 and 6 show measured values of F vs. t. On each plot we give three theoretical curves displaced by various amounts corresponding to three values of D. The solid curves represent the best values of D and the dashed curves the estimated bounds of experimental error.

The best value of D and the experimental uncertainty for each temperature is shown in Fig. 7 plotted on a logarithmic scale against the reciprocal of the absolute temperature. The measured values fall along a reasonably straight line indicating that D is given by an expression of the form

$$D=A \exp(-Q/RT). \qquad (5)$$

The values of A and Q for the straight line of Fig. 7 are given in the first row of Table I. A similar analysis of the data of Rollin gives the values in the third row.

Using some simplifying assumptions Langmuir and Dushman[7] have obtained a theoretical equation for D which relates A and Q:

$$D=\frac{Q}{Nh}\delta^2 \exp(-Q/RT), \qquad (6)$$

where $N=6.06\times10^{23}$, h is Planck's constant and δ is the lattice constant. Fitting this equation to the value of D at 950°C we get the values for A and Q in the second row of Table I.

For comparison, on Fig. 7 we have shown values for the diffusion of small concentrations of other metals into Cu. These curves were obtained by Rhines and Mehl[8] who measured the rates of diffusion for various compositions and extrapolated their results to zero concentration. Since these curves are so nearly alike, Rhines and

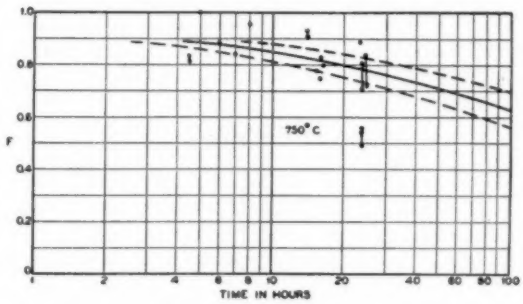


FIG. 4. Apparent activity *versus* time of heating at 750°C.

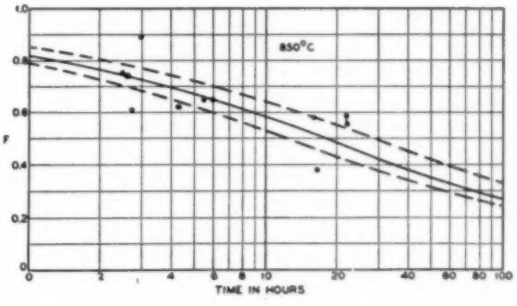


FIG. 5. Apparent activity *versus* time of heating at 850°C.

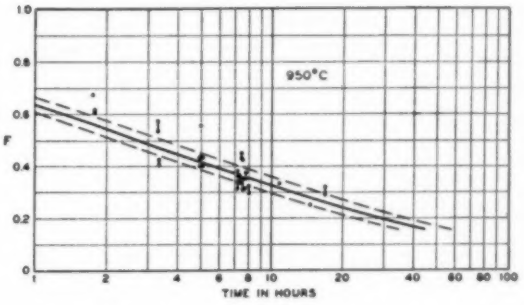


FIG. 6. Apparent activity *versus* time of heating at 950°C.

[7] S. Dushman and I. Langmuir, Phys. Rev. **20**, 113 (1922).

[8] F. R. Rhines and R. F. Mehl, Trans. A. I. M. M. E. **128**, 185 (1938).

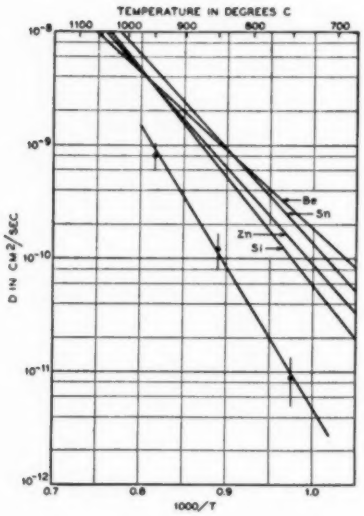


FIG. 7. Values of D *versus* $1/T$°K.

Mehl suggested that they represented something typical of copper alone and therefore the self-diffusion constant of copper. From our work we find that the self diffusion constant is smaller than the diffusion constants of the various solutes and as will be illustrated later this seems to be generally true.

In Table I we have assembled the data at present available for self-diffusion. Since diffusion involves the escape of atoms from their proper lattice sites, the energy of activation should be related to the energy required to destroy the lattice. Both the binding energy of the metal and its melting point are measures of this energy. For Cu, Au and Pb we see that Q stands in a nearly constant relationship with both these quantities and the present uncertainty in the experimental data makes it impossible to say to which Q is more closely related. Since Bi is a rhombohedral crystal, the diffusion equations involve two values for D, one parallel and one perpendicular to the C axis. The extremely large values of Q and A for the perpendicular direction are very difficult to understand. For almost any theory of diffusion, except that of Bernal[9] which involves a sequence of displacements along slip planes, the constant A should involve the product of a frequency of atomic vibration ν times the square of the lattice constant or the distance between nearest neighboring atoms δ.

[9] J. D. Bernal, Trans. Faraday Soc. **34**, 834 (1938).

For Cu, ν and δ give a value for A of ~ 1 cm²/sec.; for other elements these quantities may change by a factor of two or three producing at most changes in A of about one order of magnitude. However, a value such as 10^{46} cannot conceivably be explained in this way.

The values of Q for the self-diffusion of Cu, Ag and Pd are not inconsistent with any information now available about the three processes of diffusion usually considered; that is (1) interchange of position between pairs or groups of atoms (2) interstitial diffusion (3) diffusion by vacant lattice sites or holes. Owing to the relative incompressibility of atoms it is hard to see how they can find room to interchange positions or to become interstitial. Process (3) which does not involve the compressing of atoms but rather the creation of vacant sites seems therefore the most sound from an energetic point of view. According to this theory the diffusion constant is given by the equation

$$D = \nu\delta^2 e^{-J/RT} f, \tag{7}$$

where f is the fraction of sites which are vacant and J is the activation energy for a transition to a vacant site. If the energy required to produce a hole is H, then $f = \exp(-H/RT)$ and the activation energy for D is $H+J$. Now it is reasonable to suppose that the energy H is of the order of magnitude of the binding energy and that J is considerably smaller. (This is the case for the chemically different system Ag Cl for which the constants involved in the hole diffusion of Ag

TABLE I. *Summary of data at present available for self-diffusion.*

METAL	A CM²/SEC.	Q KCAL./MOLE	T_M MELTING POINT °K	E BINDING ENERGY KCAL./MOLE	$\frac{Q}{Tm}$	$\frac{Q}{E}$	REFERENCE
Cu	11	57.2	1356	81.2	42	0.70	α
"	0.9	51.0	"	"	38	.63	β, 8
"	47	61.4	"	"	45	.76	α, 17
Au	2.1	51	1336	92	38	.55	α, 19
"	410	62.9	"	"	47	.68	α, 18
Pb	6.66	28.05	600	47.5	47	.59	α, 12
Bi $\parallel$c	10^{-3}	31	544	47.8	57	.65	α, 20
$\perp$c	2.4×10^{46}	140	"	"	257	2.92	α, 20

α. A and Q determined from plot of log D *vs.* $1/T$°K.
β. A and Q determined from Langmuir-Dushman equation.

TABLE II. *Values of Q and of A for various metallic solutions.*

SOLVENT	GROUP	SOLUTE	GROUP	A CM²/SEC.	Q KCAL./MOLE	ATOM DIAM. SOLUTE / ATOM DIAM. SOLVENT	MELTING POINT SOLUTE °C	REFERENCE
Cu	*IB*	Cu	*IB*	11	57	1.00	1083	α
		Sn	*IVB*	0.14	45	1.21 (1.24)	231	
		Si	*IVB*	0.006	40	0.90 (1.05)	1420	
		Zn	*IIB*	0.003	38	1.08	419	β, 8
		Al	*IIIA*	0.003	38	1.10	660.2	
		Be	*IIA*	5×10^{-5}	28	0.88	1350	
Ag	*IB*	Au	*IB*	9.60	26.6	1.00	1063	
		Cu	*IB*	5.13	24.8	0.89	1083	
		In	*IIIB*	6.30	24.4	1.09	155	
		Cd	*IIB*	4.18	22.35	1.06	320.9	α, 11
		Sb	*VB*	4.58	21.7	1.12	630	
		Sn	*IVB*	6.75	21.4	1.13	231	
Pb	*IVB*	Pb	*IVB*	6.66	28.05	1.00	327.5	α, 12
		Ni	*VIIIA*	0.67	25.3	0.71	1425	γ, 13
		Sn	*IVB*	0.66	25	0.91	231	γ, 14
		In	*IIIB*	0.61	23.2	0.90	155	γ, 16
		Tl	*IIIB*	0.58	22	0.98	303.5	γ, 14
		Bi	*VB*	0.58	22	1.04	271	γ, 14
		Hg	*IIB*	0.51	19.5	0.89	−38.87	γ, 16
		Mg	*IIA*	0.49	18.7	0.92	651	γ, 16
		Ag	*IB*	0.41	15.7	0.82	960.5	γ, 14
		Au	*IB*	0.35	13.3	0.82	1063	γ, 15

α. A and Q determined from plot of log D *vs.* $1/T$°K.
β. Values of Q taken from reference 8, Fig. 14 and values of A computed from these and Fig. 15.
γ. A and Q determined from Langmuir-Dushman equation.

have been found.)[10] In connection with these ideas, however, the values of Q given in Table II are pertinent. We see that for a given solvent metal Q is greatest for self-diffusion but may decrease to about one-half that value for other chemical elements. If hole diffusion is the dominant process, the energy H should be the same for a given solvent (the concentration of the solute in all cases is small) and therefore all changes in Q should come from J. We are, therefore, forced to assume either that J is at least as large as H and is very sensitive to the nature of the solute or that another process of diffusion is involved. It would seem that calculations of the energies involved in the three theories would be very valuable.

As has been pointed out by many writers, a general trend is noticeable in the values of Q in Table II. This is summed up by saying that for a given solvent metal the activation energy is higher (that is, the diffusion is less) the more nearly the solute is like the solvent. From the tabulation, however, we see that this rule is true only in a very general way and that there is no definite criterion for similarity that is satisfactory. For Si and Sn in Cu two values are given for the ratios of atomic diameters of solute and solvent. The first of these is obtained from the crystal structure of the element and the second, in parentheses, by estimating the "effective size" of the solute atoms in dilute solution in Cu by extrapolating lattice constant *vs.* concentration plots for these alloys*· † to 100 percent solute.

Aid from the Research Corporation to Columbia University in this research is gratefully acknowledged.

APPENDIX

Let us suppose that one unit of isotope is divided between the faces of two quadrants, the individual amounts being, say, a_1 and a_2. Denoting by x_1 and x_2 the depths in the two quadrants and by C_1 and C_2 the concentrations in the sense that C_1dx_1 units of isotope lie between x_1

[10] E. Koch and C. Wagner, Zeits. f. physik. Chemie **38**, 295 (1937).

* Sn in Cu, R. F. Mehl and C. S. Barrett, Trans. A.I.M.M.E. **89**, 203 (1930).
† Si in Cu, H. F. Kaiser and C. S. Barrett, Trans. A.I.M.M.E. **93**, 1697 (1931).

and x_1+dx_1, we have the two diffusion equations

$$\partial C_1/\partial t = D\partial^2 C_1/\partial x_1^2 \quad \text{(A1)}$$

$$\partial C_2/\partial t = D\partial^2 C_2/\partial x_2^2. \quad \text{(A2)}$$

These equations are true even when the isotope is not spread uniformly in the plane of the surface.

Furthermore if there is evaporation from quadrant 1 to quadrant 2 at a rate R, we must have

$$R = D(\partial C_1/\partial x_1)_{x_1=0} = -D(\partial C_2/\partial x_2)_{x_2=0}. \quad \text{(A3)}$$

From the experiments described in II, we know that the apparent activity of material lying at a depth y from the surface is reduced by a factor $\exp(-\mu y)$. Hence the combined activity of the two quadrants is

$$F = \int_0^w e^{-\mu x_1} C_1(x_1)dx_1 + \int_0^w e^{-\mu x_2} C_2(x_2)dx_2 = \int_0^w e^{-\mu x}[C_1(x)+C_2(x)]dx, \quad \text{(A4)}$$

where w is the thickness of the disk. We see that the essential quantity is $C(x)=C_1(x)+C_2(x)$ and by adding (A1) and (A2) and the parts of (A3) we find

$$\partial C/\partial t = D\partial^2 C/\partial x^2 \quad \text{(A5)}$$

$$\partial C/\partial x = 0 \quad \text{at } x=0, \quad \text{(A6)}$$

where C and x may be considered to belong to a single quadrant which does not evaporate.

Calculation of the distribution

We must therefore solve Eq. (A5) subject to the boundary condition (A6) and to a similar boundary condition at $x=w$ and the initial condition that at $t=0$

$$C=0 \quad \text{for} \quad x\neq 0 \quad \text{and} \quad \int_0^w C dx = 1. \quad \text{(A7)}$$

The solution of the equation can be found by a method of images: one imagines an unbounded slab of material with two units of active isotope on each of the planes $x=0$, $\pm 2w$, $\pm 4w$, $\cdots$. Each of these sources spreads out in the same fashion; the concentration due to the one at $x=2nw$ is

$$u_n = (\pi Dt)^{-1} \exp(-(x-2nw)^2/4Dt). \quad \text{(A8)}$$

The concentration due to such an assembly of sources satisfies the boundary conditions and Eqs. (A5) and (A7). Therefore one form for the solution is

$$C = U = \sum_{n=-\infty}^{+\infty} u_n. \quad \text{(A9)}$$

This series converges very rapidly if Dt is not much larger than w^2.

For Dt much larger than w^2, a Fourier series solution is more rapidly convergent. The appropriate series is

$$C = V = \sum_{n=0}^{\infty} v_n, \quad \text{(A10)}$$

where

$$v_0 = 1/w \quad \text{(A11)}$$

$$v_n = (2/w)\cos(\pi nx/w)\exp(-\pi^2 n^2 Dt/w^2). \quad \text{(A12)}$$

Calculation of F and B

The above series can be used term by term to calculate F and B according to Eqs. (A4) and the corresponding equation for B given below. We find the results tabulated below:

F with U series:

$$F(z,y) = \int_0^w e^{-\mu x} U dx = \sum_{n=-\infty}^{+\infty} f_{un}, \quad \text{(A12)}$$

where

$$f_{un} = \int_0^w e^{-\mu x} u_n dx = e^{-2ny+z}\left[P\left(\sqrt{z}-n\frac{y}{\sqrt{z}}+\frac{1}{2}\frac{y}{\sqrt{z}}\right) - P\left(\sqrt{z}-n\frac{y}{\sqrt{z}}\right)\right] \quad \text{(A13)}$$

and P is defined in Eq. (4).

F with V series

$$F(z,y) = \sum_{n=0}^{\infty} f_{vn}, \quad \text{(A14)}$$

$$f_{v0} = (1-e^{-y})/y, \quad \text{(A15)}$$

$$f_{vn} = \frac{2y(1-(-)^n e^{-y})}{y^2+\pi^2 n^2} e^{-\pi^2 n^2 z/y^2}. \quad \text{(A16)}$$

B using U series

$$B(z, y)=\int_0^w e^{-\mu(w-x)}Udx=\sum_{n=-\infty}^{+\infty} b_{un}, \quad \text{(A17)}$$

$$b_{un}=e^{(2n-1)y+z}\left[P\left(-\sqrt{z}-n\frac{y}{\sqrt{z}}+\frac{1}{2}\frac{y}{\sqrt{z}}\right)-P\left(-\sqrt{z}-n\frac{y}{\sqrt{z}}\right)\right]. \quad \text{(A18)}$$

B using V series

$$B(z, y)=\sum_{n=0}^{\infty} b_{vn}, \quad \text{(A19)}$$

$$b_{v0}=(1-e^{-y})/y, \quad \text{(A20)}$$

$$b_{vn}=\frac{2y((-1)^n-e^{-y})}{y^2+\pi^2n^2}e^{-\pi^2n^2z/y^2}. \quad \text{(A21)}$$

These series have been used in calculating the curves of Figs. 2 and 3; the U series used for $z<y^2$ and the V series for $z>y^2/9$ the region of overlap serves as a check on the calculations.

References of Tables

[11] W. Seith and E. A. Peretti, Zeits. f. Elektrochem. **42**, 570 (1936).
[12] G. v. Hevesy, W. Seith and A. Keil, Zeits. f. Physik **79**, 197 (1932).
[13] W. Seith, E. Hofer and H. Etzold, Zeits. f. Elektrochem. **40**, 322 (1934).
[14] W. Seith and J. G. Laird, Zeits. f. Metallkde. **24**, 193 (1932).
[15] W. Seith and H. Etzold, Zeits. f. Elektrochem. **40**, 829 (1934).
[16] W. Seith, Zeits. f. Elektrochem. **41**, 872 (1935).
[17] B. V. Rollin, Phys. Rev. **55**, 231 (1939).
[18] A. Sagrubskij, Physik. Zeits. Sowjetunion **12**, 118 (1937).
[19] H. A. C. McKay, Trans. Faraday Soc. **34**, 845 (1938).
[20] W. Seith, Zeits. f. Elektrochem. **39**, 538 (1933).

JULY 1, 1939 PHYSICAL REVIEW VOLUME 56

Natural Radioactivity of Lutecium

W. F. Libby
Department of Chemistry, University of California, Berkeley, California

(Received April 22, 1939)

The half-life of the natural radioactivity of lutecium reported by Heyden and Wefelmeier is $2.9\pm0.5\times10^{12}$ years for the element or $7.3\pm2\times10^{10}$ years for Lu^{176}, which probably is the active isotope. Negative electrons are emitted. The upper energy limit of the spectrum is 215,000±15,000 electron volts.

THE bearing of the natural radioactivities of the ordinary elements on nuclear physics and geophysics makes it desirable to have data on their characteristics which are as accurate as those for the shorter-lived artificial radioactivities. It is the purpose of this paper to present data obtained on the natural radioactivity of lutecium (cassiopeium) reported by Heyden and Wefelmeier.[1]

The work of Gollnow[2] on the hyperfine structure of lutecium had shown the presence of an unreported isotope of about 2.5 percent abundance. Recently Mattauch and Lichtblau[3] investigated the element with a mass spectrograph and found Lu^{176} present to 2.5±0.07 percent of the element. Since both ytterbium and hafnium have isotopes of mass 176, this makes the Lu^{176} case somewhat similar to that of K^{40} and makes a determination of the sign of the beta-radiation of considerable interest.

The lutecium used was kindly furnished by Professor B. S. Hopkins of the University of Illinois. It was pure except for traces of thulium and ytterbium.

Proof That the Activity Belongs to a Rare Earth Element

A consideration of the possible impurities (other than elements of low atomic number which would have to be present in large amounts to affect the data) reveals that all of those emitting beta-radiation of approximately the same

[1] M. Heyden and W. Wefelmeier, Naturwiss. **26**, 612 (1938).
[2] H. Gollnow, Zeits. f. Physik **103**, 443 (1936).
[3] J. Mattauch and H. Lichtblau, Zeits. f. Physik **111**, 514 (1939).

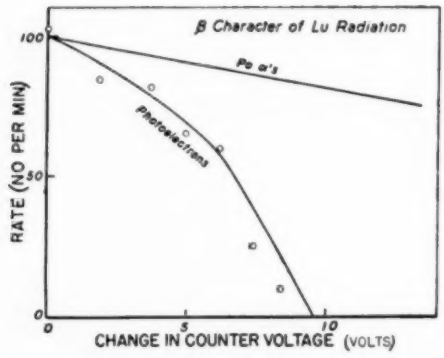


FIG. 1. Data obtained with proportional counter.

energy as that from lutecium either will come to equilibrium with alpha-emitting products or will be precipitated with PbS. The activity of the original Lu_2O_3 was compared with that of the purified compound and found to differ less than five percent. The purification was conducted by dissolving the Lu_2O_3 in dilute HNO_3, adding $Pb(NO_3)_2$, and saturating the solution with H_2S. The Lu_2O_3 was prepared from the filtrate from the PbS separation.

The search for alpha-radiation was made by using a screen wall counter[4, 5] with a wire 0.22 millimeter in diameter which allowed proportional counting to be done by lowering the voltage applied across the counter. The points in Fig. 1 show the data obtained. The ordinates are the counts observed when a Lu_2O_3 sample is placed over the counter. The abscissae are voltages below a standard value at which both alpha- and beta-radiation counted. The upper curve for alpha-radiation was obtained by means of a polonium source and a mica window in the counter chamber and the lower curve shows the effect due to gamma-radiation from a thorium source. This is characteristic of beta-radiation, since the gamma-radiation ejects photoelectrons which count. Evidently any alpha-radiation present must be less than about ten percent of the total radiation. This fact, together with the PbS test shows that the activity must belong to a member of the rare earth group.

PROOF THAT THE ELECTRONS ARE NEGATIVELY CHARGED

Flat vanes running parallel to the counter axis and inclined so their surfaces made an angle of about 30° with the plane through the center of the vane and the counter wire (cf. reference 4) were inserted between the sample cylinder and screen wall counter. This device allows the sign of the beta-radiation to be determined by observation of the variation of the counting rate when the direction of the magnetic field is changed. Fig. 2 presents the results. The upper curve is drawn through the data obtained when the field direction was chosen to help negative electrons pass through the vanes and the lower curve is that obtained when the field was reversed to pass

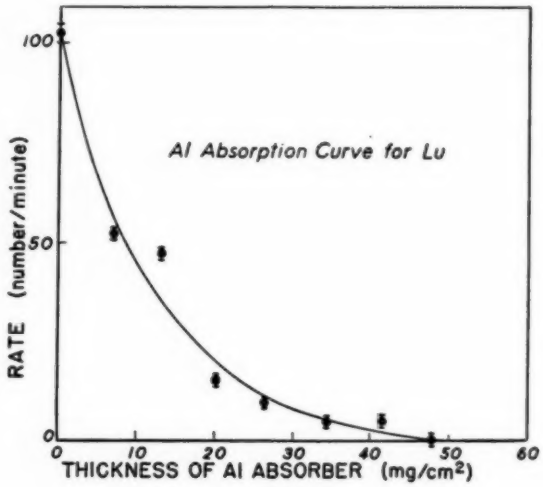


FIG. 3. Aluminum absorption curve.

[4] W. F. Libby, Phys. Rev. **46**, 196 (1934).
[5] W. F. Libby, Phys. Rev. **55**, 245 (1939).

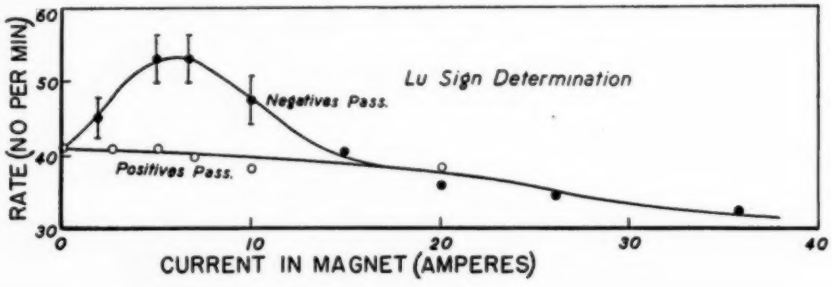


FIG. 2. Data on sign of beta-radiations.

positive electrons. The results show that the radiation is predominantly negative in character, the uncertainty amounting to a possible twenty percent. The conclusion is that the process occurring is

$$Lu^{176} \rightarrow \beta + Hf^{176}.$$

Energy of Radiation

Figure 3 presents data on the aluminum absorption curve for the radiation. The Lu_2O_3 sample was placed around the screen wall counter and various thicknesses of foil were placed around the screen wall. Correction was made for the absorption by the gas in the counter (0.20 mg/cm²). Apparently the end-point is at 50 ± 5 mg/cm² of aluminum. From the data of Eddy[6] and Varder[7] on the ranges of cathode rays of various energies in aluminum, this gives $215{,}000\pm15{,}000$ electron volts as the upper energy limit of the lutecium spectrum. This procedure seems to be the most reliable method for interpreting absorption data on beta-spectra with upper limits below about 700,000 ev (cf. reference 5).

A determination of the upper limit by use of the screen wall counter in a magnetic field was made also. Fig. 4 gives the data representing the counting rate as a function of the current in the magnet coils. The end-point is at 50 ± 5 amperes which corresponds to 2000 ± 200 gauss and gives 1725 ± 173 as the $H\rho$ value for the upper limit of the beta-spectrum (cf. reference 5). This sets $212{,}000\pm18{,}000$ ev as the upper energy limit. Since the magnetic and absorption methods agreed, the end-point appears to be $215{,}000\pm15{,}000$ ev. The value agrees well with the value

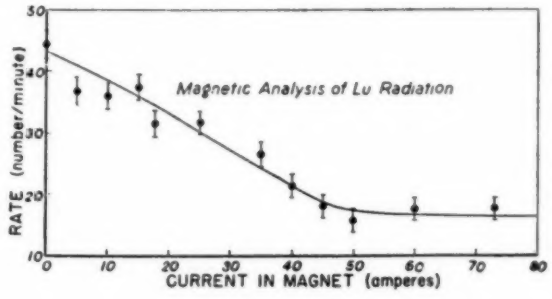


Fig. 4. Magnetic data.

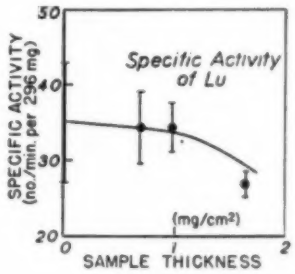


Fig. 5. Data on specific activity.

of 45μ of Al for the absorption half-thickness given by Heyden and Wefelmeier.[1]

Half-Life Determination

The specific activity was determined by using thin layers of Lu_2O_3 (obtained by moistening with ethyl alcohol, spreading smoothly on the sample cylinder, and evaporating the alcohol). The number of counts per minute per unit weight of Lu_2O_3 was plotted against the average thickness of the samples. Fig. 5 shows the data. A curve of the same shape as that found experimentally for potassium and rubidium[8] (corrected for differences in the penetrating powers of the radiations) was drawn through these points and used as the specific activity without self-absorption. The factor for geometrical arrangement and counting efficiency was determined by substituting potassium chloride and calculating from the known weight of potassium used and the known decay constant. The result (8.3 percent counted) agreed within the limits of measurement with that calculated for the geometrical factor alone, showing that the counting efficiency was near one hundred percent. The specific activity of 34 ± 6 counts per minute per 296 mg of Lu_2O_3 obtained gives

$$34\times(199/0.296)\times(1/0.083)=2.75\times10^5$$

per mole per minute as the activity of lutecium element. Since Lu^{176} (2.5 percent) is probably the isotope responsible this corresponds to 1.1×10^7 per minute per mole of Lu^{176} which gives $7.3\pm2\times10^{10}$ years as the half-life of Lu^{176}. This figure agrees well enough with that of 4×10^{10} years given by Heyden and Wefelmeier.

The author wishes to thank Professor B. S. Hopkins for the Lu_2O_3 which he so generously furnished for this research.

[6] C. E. Eddy, Proc. Camb. Phil. Soc. **25**, 50 (1929).
[7] R. W. Varder, Phil. Mag. **29**, 726 (1915).
[8] W. F. Libby, Phys. Rev. **46**, 745 (1934).

A Precise Determination of the Energy Released in the Production of Deuterium from Beryllium under Proton Bombardment

LESTER S. SKAGGS
University of Chicago, Chicago, Illinois
(Received April 20, 1939)

The kinetic energy of the deuterons liberated in the reaction Be^9 (p, d) Be^8 has been measured by deflecting the deuterons in an accurately known electrostatic field. The technique of the use of the electrostatic deflector is discussed, and experimental arrangements appropriate to the determination are described. From measurements at proton energies of 258, 262, and 315 kv the energy released in the reaction is found to be 0.557 ± 0.006 Mev. The shift of the high energy limit of the deuterons with bombarding voltage establishes the masses of the particles involved. By combining the result with other energy data as known at present it is shown that (*a*) the threshold for the disintegration of Be^9 by gamma-rays is 1.62 ± 0.02 Mev. (*b*) Be^8 is stable with respect to disintegration into two alpha-particles by 0.174 ± 0.09 Mev. Revision of the present values of the mass-spectrographic bracket between deuterium molecules and helium, or the value of the energy released in the reaction Li^6 $(d, \alpha)He^4$ would necessitate revision of conclusion (*b*). The experimental evidence on the stability of Be^8 is briefly discussed.

INTRODUCTION

IN a previous communication[1] from this laboratory we have described an electrostatic analyzer, adapted to the precise measurement of the kinetic energies of slow particles released in certain nuclear reactions. In this first communication we reported the measurement of the energy of the alpha-particles emitted in the reaction Be^9 (p, α) Li^6. In a subsequent brief letter,[2] we announced the results of the determination of the deuteron energy in the concomitant reaction Be^9 (p, d) Be^8. The present paper will give some experimental details of this latter investigation and discuss the results.

EXPERIMENTAL PROCEDURE

No change in the electrostatic analyzer itself was made in the interim between the alpha-particle and the deuteron experiments, and the previous description[1] is adequate. The main features may be summarized as follows. Some of the disintegration particles, produced from a polished target of beryllium under proton bombardment, pass through an entrance slit 0.1 cm wide placed 2.3 cm from the target. This slit is 8.90 cm (the focal distance) from the edges of the cylindrical plates of the deflecting electrostatic field. These plates are spaced 0.635 cm apart and have an average radius of curvature of 25.400 cm. They subtend an angle of 90° at their center of curvature, and are 8.26 cm high, although the height of the beam going through them is only 1.5 cm.

The electrostatic analyzer was disassembled shortly after these experiments were completed and the dimensions of the plates re-measured. The separation of the plates at no point varied by more than one percent of the mean 0.635 cm separation. These small areas of variation occurred at isolated points along the plates, mostly near the edges and not in the path of the beam of particles being analyzed. They therefore introduced no appreciable error.

The arrangement of the ionization chamber in which the particles were counted after passing through the analyzer is shown in Fig. 1. The entrance slit of the chamber is 1.5 cm high and 0.3 cm wide. It is covered with perforated screen with holes 0.0216 cm in diameter, and has a transmission ratio of 0.321.[3] A collodion foil of 2 mm stopping power is placed on the inside of the ionization chamber over the screen. The parts of the foil not actually over the window are painted with colloidal carbon (Aquadag) to make a conducting surface. The distance from the collecting electrode to the screen of the chamber was 0.8 cm and a potential of 750 volts was applied to collect the ions. Since the ranges of both types of particles emitted from beryllium

[1] S. K. Allison, L. S. Skaggs and N. M. Smith, Jr., Phys. Rev. **54**, 171 (1938).
[2] S. K. Allison, E. R. Graves, L. S. Skaggs and N. M. Smith, Jr., Phys. Rev. **55**, 107 (1939).
[3] R. B. Bowersox, Phys. Rev. **55**, 323 (1939).

under proton bombardment are about 0.8 cm in air, this means that the paths actually ended in the air of the ionization chamber. Under these conditions the amplified pulses from the particles entering the chamber were shown to be about three-fourths the maximum height of alpha-particle pulses obtainable from a polonium source shooting particles into the same chamber. It was not possible, however, to make a satisfactory distinction between alpha-particle and deuteron pulses from their relative heights as has been done by other observers[4] in shallower chambers. The deuteron and alpha-particle limits were identified by another method, to be described later.

The deflecting voltages put on the plates of the analyzer ranged up to 38 kilovolts for the beryllium particles. The rectified high voltage current from a transformer-kenetron circuit with half-wave rectification was passed into a condenser of 0.1 microfarad capacity, and a metallized resistance of 2.9×10^8 ohms placed between this condenser and the electrostatic deflector, which was also directly connected to an electrostatic voltmeter of sensitivity 300 volts per scale division in the region 25–45 kilovolts. No appreciable current was drawn from the condenser during the deflection of the particles through the analyzer. After a run, the voltmeter scale was calibrated in several places by connecting the voltmeter directly to the condenser, and putting in parallel with it a 2.8×10^7 ohm wire-wound precision resistor which at 38 kv drew 1.4 ma from the condenser. The voltage drop when this current passed through a resistance of approximately 1000 ohms was measured in a potentiometer circuit. The constants in this circuit were recently re-checked by Mr. F. V. Stearns, and found accurate to 0.1 percent.

As explained in a previous paper,[1] the geometry of the electrostatic analyzer is such that if V is the deflecting voltage which causes a beam of particles of homogeneous energy to traverse the mean radius and give the maximum counting rate in the ionization chamber, a change dV will reduce the counts to zero, where $dV/V=0.8$ percent. This is a measure of the resolving power. From the voltage V at which a homogeneous particle group traverses the mean radius, the energy E in electron volts of the particles in the group may be computed from

$$E=20.00Vz,$$

where z is the charge on the particles in electron units.

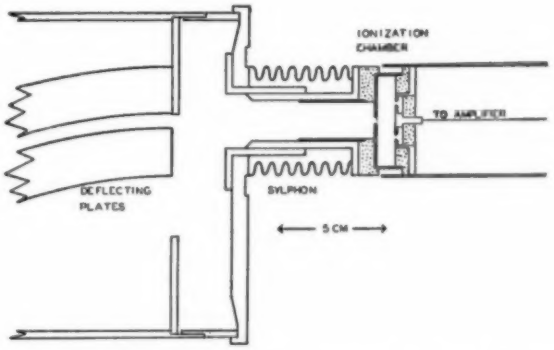


FIG. 1. Arrangement of the ionization chamber for detection of particles which have passed through the electrostatic analyzer. The front cap of the ionization chamber serves as part of the vacuum wall.

The target was a disk of pure beryllium metal 1.6 cm in diameter, mounted and heated previous to each bombardment as described by Hatch.[4] The proton beam was limited to a diameter of 8 mm by circular apertures in front of the target. The magnetically analyzed proton beam currents were from one to two microamperes in intensity. The plane of the polished target face was at 45° to the proton beam, and the disintegration particles entering the analyzer made an angle of 90° with the proton beam.

METHOD OF TAKING READINGS

The deflecting voltage was varied in steps of approximately 1000 volts through the region in which it was expected that the particles would traverse the proper radii through the analyzer. A cathode-ray oscilloscope was always connected to the output of the linear amplifier. If the counts did not exceed 2 to 3 per second, visual counting by two observers watching the oscilloscope was utilized. Frequent check counts with the proton beam off the target were made, especially at low counting rates. For higher counting rates (the 315- and 351-kv curves of

[4] M. L. E. Oliphant, A. E. Kempton and Lord Rutherford, Proc. Roy. Soc. Lond. **A150**, 241 (1935); J. S. Allen, Phys. Rev. **51**, 182 (1937); F. Kirchner and H. Neuert, Physik. Zeits. **38**, 969 (1937); G. T. Hatch, Phys. Rev. **54**, 165 (1938).

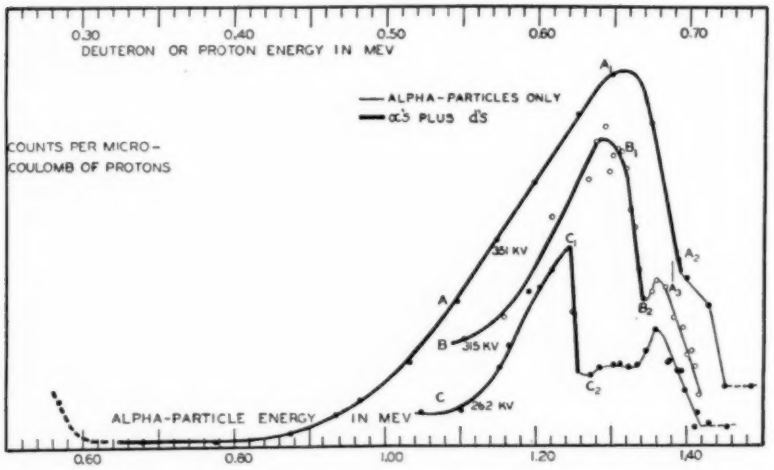


FIG. 2. The energy spectra of the particles emitted from beryllium under proton bombardment. For absolute values of the ordinates see Table I in the text.

Fig. 2), a scale-of-sixteen circuit with a mechanical counter was used, while an observer watched the oscilloscope for spurious disturbances. The elimination of such spurious counts has been one of the most difficult features of the work. Proton beam currents and readings of the current through the 7.5×10^9-ohm resistor measuring the bombarding voltage were taken at frequent intervals.

Discussion of Results

Figure 2 shows curves obtained at different bombarding energies. The total number of particles counted, the rate of counting and the method of counting (scale-of-sixteen with mechanical counter or visual observation of oscilloscope) are given in Table I for various points on the curves. Curve C of the figure shows the deuteron and alpha-particle limits well separated when the bombarding voltage was 262 kv. But when the bombarding voltage was raised to 351 kv the limits became practically coincident as is shown by curve A. It can be shown in the following way that the limit on the left of curve C is due to deuterons. The ordinary laws of mechanics give the following relation for the energy of the deuterons, ejected at 90° to the proton beam:

$$E_d=0.7E_p+0.8Q_1,$$

where E_p is the energy of the bombarding protons and Q_1 is the reaction energy. For the energy of the alpha-particles, the corresponding relation is:

$$E_\alpha=0.5E_p+0.6Q_2,$$

where Q_2 is the energy release in the reaction Be^9 (p, α) Li^6. If the bombarding energy is increased by 89 kv then the energy of the deuterons should be increased by 62 kv and the energy of the alpha-particles by 45 kv. If we assign the peak on the left of curve C to deuterons, then it indicates that the deuterons have an energy of 0.628 Mev at 262 kv bombarding energy. When the bombarding energy is raised to 351 kv the energy of the deuterons should be increased by 0.062 Mev. They should then have an energy of 0.690 Mev and the position of the limit at 351 kv bombarding energy should be as is shown by the point A_3 on curve A. If this were wrong and the limit on the left due to alpha-particles, their energies would be 1.256 Mev, which would be shifted to 1.301 Mev at 351 kv bombarding energy. The form of curve A clearly indicates that the original assumption of deuterons was correct. The shift of the limit on the high energy end of the entire curve is also consistent with the assumption that it is due to alpha-particles, the curves indicating a shift of about 40 kv and the calculations requiring 45 kv. Furthermore curve B, taken at an intermediate voltage, indicates the shifting of the limits to the right as the bombarding voltage is increased and the positions of

its limits are in agreement with the conclusion just reached.

Because of the better resolution of the two limits when the bombarding energy is somewhat lower than 350 kv, we shall obtain our value of Q_1 entirely from curves taken in that region of bombarding energy. Table II shows the results of five trials. The correction of -0.005 Mev applied between columns 4 and 3 of Table II is arrived at as previously explained.[1] Due to the width of the beam of particles (0.1 cm) at the window of the ionization chamber, and the width of this window itself (0.3 cm), a homogeneous beam of particles will just reach zero counting rate when the deflecting voltage is raised 0.8 percent above that at which the particles traverse the mean radius. Thus a correction of -0.8 percent is applied to the high energy edges of the alpha-particle and deuteron groups.

In the vicinity of point B_1, the solid curve is drawn in a rather arbitrary manner through scattered points. If we follow the points consecutively, a curve is obtained which shows a fine structure of the deuteron energies. We have spent considerable time trying to establish the reality of such a fine structure. It appears on several curves, and is larger than the expected statistical error, but it does not appear with equal prominence on all the curves. At the present stage of our experiments, we do not wish to state categorically that such a fine structure is actually present. We are continuing investigation of this energy region. If such a structure is present, it would probably represent a very low excited state of Be^8, about 16 kv above the ground state. Another irregularity in the curve indicates such a state at 30–50 kv.

The formula used for calculating the reaction energy assumes an angle of 90° between the incident particle and the ejected particle. It can be shown that the variation in the value of Q with the angle θ between the incident and ejected particles is given by

$$dQ=2\left(\frac{M_1M_2}{M_3}E_1E_2\right)^{\frac{1}{2}}\sin\theta d\theta,$$

where the subscripts 1, 2 and 3 refer to the incident, the ejected and the residual nuclei, respectively, M is the mass of the nucleus and E its kinetic energy. The angle between the direction of the incident beam and the tangent to the circular deflection plates of the analyzer at the entrance of the field was measured and found to be $88°\,40'\pm20'$. To this should be added the maximum allowable angle in the forward direction between the tangent to the deflection plates and the direction of a particle which can just pass through the analyzer. This is about $1°\,10'$. Therefore the most energetic particles which come through the analyzer were ejected at an angle of $89°\,50'\pm20'$ with the bombarding protons. Since the measurements are made on the most energetic particles no appreciable error results from the assumption made in the calculations of a right angle between the two beams.

In the case of curve A of Fig. 2, the deflecting voltage on the electrostatic analyzer was lowered to the point where scattered protons could traverse the required path. The energies of such protons can be obtained from the equation[5]

$$E_s/E_p=\frac{[M\cos\theta\pm(m^2-M^2\sin^2\theta)^{\frac{1}{2}}]^2}{(M+m)^2},$$

where E_p is the bombarding energy, M is the proton mass, and m the beryllium mass. Taking θ as 90° we find $E_s=280$ kv if $E_p=351$ kv. No care was taken in observing these pulses but

TABLE I. *Number of particles counted in curves of Fig. 2.*

POINT	TOTAL NO.	NO. PER SEC.	METHOD OF COUNTING
A_1	744	18.6	Mechanical
A_2	400	10.0	Mechanical
B_1	675	11.2	Mechanical
B_2	270	4.5	Mechanical
C_1	246	2.36	Visual
C_2	90	1.05	Visual

TABLE II. *Energy liberated in the reaction* $Be^9(p, d)Be^8$.

TRIAL	PROTON ENERGY KV	DEUTERON LIMIT MEV	DEUTERON LIMIT CORRECTED	Q_1 MEV	Q_1 MILLI-MASS-UNITS
1	258	0.629	0.624	0.555	0.596
2	262	0.628	0.623	0.557	0.598
3	262	0.628	0.623	0.557	0.598
4	315	0.670	0.665	0.556	0.597
5	315	0.673	0.668	0.559	0.600

[5] C. G. Darwin, Phil. Mag. **27**, 499 (1914).

they were obviously present at the proper deflecting voltage of 14 kv.

Inspection of the relative heights of the limits in the 262-kv curve shows that the numbers of deuterons and alpha-particles are approximately equal. This is in agreement with observations of others by counting at different biases on the Thyratron scaling circuit and thus distinguishing between the intense alpha-particle pulses and the relatively weaker deuteron pulses.[4]

In view of the results of Table II, and of the previous paper from this laboratory,[1] we can now announce the energy releases in the following reactions:

$$\mathrm{Be^9+H^1=Be^8+D^2}+Q_1; \quad Q_1=0.557\pm0.006 \text{ Mev} \tag{1}$$

$$\mathrm{Be^9+H^1=Li^6+He^4}+Q_2; \quad Q_2=2.152\pm0.04 \text{ Mev.} \tag{2}$$

Deductions from the Energy Values

By adding the equation

$$\mathrm{D^2}+h\nu=\mathrm{H^1}+n^1+Q_4$$

to Eq. (1) above, we obtain

$$\mathrm{Be^9}+h\nu=\mathrm{Be^8}+n^1+Q_1+Q_4.$$

In a new determination of Q_4 by Rogers and Rogers,[6] the value -2.174 ± 0.02 Mev has been obtained. Less recently Stetter and Jentschke[7] found -2.189 ± 0.022 Mev. Taking the average of these as -2.18 ± 0.02 Mev, we obtain the value -1.62 ± 0.02 for the sum Q_1+Q_4. This numerical value represents the binding energy of the odd neutron in $\mathrm{Be^9}$, and also is the photoelectric threshold for the photodisintegration of $\mathrm{Be^9}$. Collins, Waldman and Polye[8] have reported that this reaction initiates with the x-rays from a high voltage tube when the voltage rises to about 1.65 Mev.

Our results also have some bearing on the problem of the stability of $\mathrm{Be^8}$. If we use, in addition to the equations (1) and (2) representing our work, the equation

$$\mathrm{Li^6+D^2=2He^4}+Q_3$$

we can obtain

$$\mathrm{Be^8-2He^4=-(2D^2-He^4)}-Q_1+Q_2+Q_3.$$

The term $(\mathrm{2D^2-He^4})$ is a well-known mass-spectrographic doublet for which Bainbridge[9] finds the value 25.61 ± 0.04 mMU (thousandths of a mass unit). The value of Q_3 has been measured quite accurately by Oliphant, Kempton and Lord Rutherford,[10] who obtained the value 22.08 ± 0.07 Mev. From these values we find that the term $\mathrm{Be^8-2He^4}$, which represents the stability of beryllium eight with respect to disintegration into two alpha-particles, to be -0.187 ± 0.10 mMU. This indicates that beryllium eight is stable by 0.174 ± 0.09 Mev.

Various other data may be used along with ours to calculate the stability of beryllium eight, although it seems[11] that the above is the most reliable. If Aston's value[12] of 25.51 ± 0.08 mMU for the deuterium-helium bracket is used instead of Bainbridge's, we get by the above scheme a $\mathrm{Be^8}$ stable with respect to two alpha-particles by 0.087 ± 0.12 mMU. This is a result indecisive as to stability or instability.

If we discard the reaction $\mathrm{Li^6}\ (d,\alpha)\ \mathrm{He^4}$ as the source of the needed additional data, and use instead

$$\mathrm{Be^9}\ (d,\alpha)\ \mathrm{Li^7}; \quad Q_5=6.95\pm0.12 \text{ mMU},[13]$$

and $$\mathrm{Li^7}\ (p,\alpha)\ \mathrm{He^4}; \quad Q_6=17.13\pm0.006 \text{ mMU},[14]$$

we obtain (after the Q's are expressed in mass rather than energy units) the result

$$\mathrm{Be^8-2He^4=-(2D^2-He^4)}-Q_1+Q_5+Q_6=-0.33\pm0.14 \text{ mMU}$$

with Bainbridge's value of the bracket. If Aston's value is used, the stability is 0.23 ± 0.16 mMU.

Thus the *present status* of the energy values definitely points to a *stable* $\mathrm{Be^8}$, with the best value of the stability 0.174 ± 0.09 Mev.

$\mathrm{Be^8}$ has not been found as a natural constituent of the earth's crust. Bleakney, Blewett,

[6] F. T. Rogers and M. M. Rogers, Phys. Rev. **55**, 263 (1939).
[7] G. Stetter and W. Jentschke, Zeits. f. Physik **110**, 214 (1939).
[8] G. B. Collins, B. Waldman and W. R. Polye, Phys. Rev. **55**, 412 (1939).
[9] K. T. Bainbridge and E. B. Jordan, Phys. Rev. **51**, 384 (1937).
[10] M. L. E. Oliphant, A. E. Kempton and Lord Rutherford, Proc. Roy. Soc. Lond. **A149**, 406 (1935); cf. S. K. Allison, Phys. Rev. **55**, 624 (1939).
[11] S. K. Allison, Phys. Rev. **55**, 624 (1939).
[12] F. W. Aston, Proc. Roy. Soc. Lond. **A163**, 391 (1937).
[13] J. H. Williams, R. O. Haxby and W. G. Shepherd, Phys. Rev. **52**, 1031 (1937).
[14] Oliphant, Kempton and Rutherford, reference 10.

Sherr and Smoluchowski[15] have examined the mass spectrum of beryllium under conditions in which doubly charged oxygen could not contribute a false mass eight spot, and found no Be^8 to one part in 10,000 of Be^9.

O. Laaff[16] placed an ionization chamber of effective depth 0.1 cm of air behind a window of 0.25 cm stopping power and amplified the pulses from alpha-particles produced in the boron-proton reactions. It is well known that there is a continuous distribution of ranges, culminating in a discrete group at 4.4 cm, supposed to arise from the reaction $B^{11}(p, \alpha)Be^8$ in which the alpha-particle recoils from a complete Be^8 nucleus. The energy released is 8.55 Mev, so that the recoil energy of the Be^8 nucleus is 2.85 Mev, disregarding the 185 kv bombarding energy used by Laaff. Occasional pulses, more intense than those from alpha-particles, were observed in the chamber, and attributed to simultaneous entrance of two alpha-particles from the disintegration of the residual Be^8. From the effective solid angle for large pulses, which appeared to be less than the geometrical solid angle of the apparatus, Laaff estimated that the energy released in the disintegration of the Be^8 into two alpha-particles was smaller than 200 kv.

It seems difficult to ascribe the presence of intense pulses to two simultaneous alpha-particles without a discussion of pulse heights. The theory of ionization by moving positive charges predicts that, at the same velocity, the loss of energy per cm of path by a moving, quadruply charged, Be^8 nucleus is four times that of a single alpha-particle, and, of course, twice that of two alpha-particles. But if the presence of the two simultaneous alpha-particles be granted, it does not seem improbable that a Be^8 nucleus, stable by only 0.174 Mev, should be thrown into a low, oscillating, virtual quantum state from which it can decompose; particularly when the available recoil energy is 2.85 Mev. The recoil energy available for the Be^8 nucleus in the reaction $Be^9(p, d)Be^8$, which we have studied here, is only 0.11 Mev.

Glückauf and Paneth[17] exposed a sample of beryllium metal to the gamma-radiation from RaB and RaC, dissolved the beryllium, and tested the evolved gases for traces of helium. They found that helium actually had been produced in the beryllium, to an amount of the order of magnitude of that to be expected from the yield of neutrons in the photodisintegration of Be^9. Thus they state that the end product of the photodisintegration is two alpha-particles and a neutron, and that Be^8, if produced in an intermediate stage of this reaction, is not stable.

The energy required to disintegrate Be^9 into two alpha-particles and a neutron according to our results is $1.62+0.17$ or 1.79 Mev. The two gamma-ray lines from RaB and RaC which have energy enough to disintegrate beryllium at all are 1.78 and 2.22 Mev. Thus they are also probably able to produce the complete disintegration into helium and neutrons.

The author wishes to thank Professor Samuel K. Allison for his interest and assistance in this work. Others who have given help are Mrs. E. R. Graves, Mr. Leonard Miller, Mr. Nicholas M. Smith, Jr., and Mr. Gilbert Perlow.

[15] W. Bleakney, J. P. Blewett, R. Sherr and R. Smoluchowski, Phys. Rev. **50**, 545 (1936).
[16] O. Laaff, Ann. d. Physik **32**, 743 (1938).
[17] E. Glückauf and F. A. Paneth, Proc. Roy. Soc. **165**, 229 (1938).

An Atomic Beam Study of the Hyperfine Structure of the Metastable $^2P_{3/2}$ State of $In^{115}I$

The Electric Quadrupole Moment of In^{115}*

DONALD R. HAMILTON†
Columbia University, New York, New York

(Received May 15, 1939)

The zero-moment method of atomic beams has been used to investigate the hyperfine structure of the metastable $^2P_{3/2}$ state of $In^{115}I$. Thirteen zero-moment peaks have been observed, and it is found that their behavior is completely described by the equation for the h. f. s. energy levels, $E=\frac{1}{2}aC+bC(C+1)$, together with a knowledge of the nuclear g factor. a and b are the interval rule and quadrupole interaction constants, respectively; the numerical values are $b/a=(6.432\pm0.04)\times10^{-3}$; $a=(8.10\pm0.10)\times10^{-3}$ cm^{-1}; $b=(0.0521\pm0.0009)\times10^{-3}$ cm^{-1}. From this the total h. f. s. separation is (0.1184 ± 0.0015) cm^{-1} and the quadrupole moment of In^{115} is 0.84×10^{-24} cm^2. The nuclear spin is verified to be 9/2. If there is an additional term $c(C^3+4C^2+4/5C)$ in the h. f. s. energy level equation, as suggested by Tolansky in the case of iodine, an upper limit of 5×10^{-6} may be set on c/a.

THE electric quadrupole moments of a number of nuclei have been measured spectroscopically in recent years by observing deviations from the interval rule in hyperfine structure (h. f. s.). The interval rule arises from the cosine interaction between the nuclear magnetic moment and the magnetic field produced at the nucleus by the electrons. The nuclear quadrupole moment, Q, is defined as the average of $(3z^2-r^2)$ over the nuclear charge density for the state $M_I=I$, i.e., $eQ=\int\rho(3z^2-r^2)d\tau$, where $\int\rho d\tau=Ze$. Q is thus a measure of the deviation of the nucleus from spherical symmetry and gives rise to a cosine squared interaction with a properly asymmetrical electronic charge density. The position of the h. f. s. levels is given by

$$E=\tfrac{1}{2}aC+bC(C+1)$$
$$(C=F(F+1)-I(I+1)-J(J+1)), \quad (1)$$

where C is the quantum-mechanical analog of the cosine. The first term gives the interval rule and the second arises from the nuclear quadrupole. b is proportional to Q and to $(3\cos^2\theta-1)/r^3$ averaged over the electronic state $M_J=J$. Perturbations by other states may also give rise to deviations from the interval rule.

With the zero-moment method of atomic beams,[1] which is used in the present experiment, one observes in a magnetic field the behavior of the h. f. s. levels of the state in which the atoms of the beam exist. If one follows the course of these levels as the field is increased from zero through the region of transition between the h. f. s. anomalous Zeeman and Paschen-Back effects, it is found that for each of a number of levels there exists a value of the field at which $dE/dH=0$. An atom for which this is true will experience no net translational force in an inhomogeneous magnetic field of this value, or in other words its effective magnetic moment is zero. If, now, an atomic beam traverses a transversely inhomogeneous deflecting field and the intensity of the beam at the position of zero deflection is observed as a function of magnetic field, "peaks" will be observed at the zero-moment fields of the various states. The fields at which such peaks occur are the observed quantities of the experiment.

Because of the limitations of the surface ionization beam detection method, set by the ionization potential of the atom being detected, zero-moment experiments have so far been limited to the alkali metals and indium. The ground states of all these elements have electronic charge distributions for which $[\langle(3\cos^2\theta-1)/r^3\rangle_{Av}]_{M_J=J}$ is zero, hence a behavior not influenced by a nuclear quadrupole.

However, in the absolute moment experiments of Millman, Rabi and Zacharias carried out on

* Publication assisted by the Ernest Kempton Adams Fund for Physical Research of Columbia University.
† University Fellow, 1938–1939.
[1] V. W. Cohen, Phys. Rev. **46**, 713 (1934).

the $^2P_{1/2}$ ground state of indium[2] (and not to be confused with the more recent magnetic resonance moment experiments), it was found possible to observe peaks of the metastable $^2P_{3/2}$ state, which lies 2212.6 cm^{-1} above the ground level. The measurements with which this paper is concerned were made on these peaks with a new apparatus which has two advantages, for this purpose, over the one used by Millman, Rabi and Zacharias: first, increased resolving power, giving more and better peaks; second, use of the well-known two-wire type deflecting magnet[3] which has no iron and thus makes it possible to avoid magnetization curve difficulties. At the source temperature of about 1500°K, the expected abundance of the metastable state is about 20 percent according to Boltzmann statistics; since this level has 40 magnetic states, each peak should have about 0.5 percent of the intensity of the entire beam. This statement is verified by the authors cited and in the experiment to be described. Since $[\langle(3\cos^2\theta-1)/r^3\rangle_{Av}]_{M_J=J}$ is not zero for the $^2P_{3/2}$ state, the fields at which the peaks occur depend on Q and a means of measuring Q is thus provided.

This occurrence, for certain states, of a value of the magnetic field at which the effective moment of an atom is zero is a result of the progressive reorientation of **J** by the external field, which uncouples **I** and **J**. In this reorientation process there is some value of the magnetic field at which the time average direction of **J** is perpendicular to **H**, and the effective moment of the atom is then zero. The strength of the field which is necessary to bring this about is directly proportional to the strength of the interaction between the electrons and the nuclear magnetic dipole moment if one neglects the influence of the quadrupole interaction. The coupling of the nuclear quadrupole to the asymmetrical electron charge distribution brings into play another torque on **J** with a different dependence on the relative orientations of **I** and **J**. This helps or hinders the rearrangement process, as the case may be, and causes the zero moment to occur at a higher or lower value of the field than in the absence of the quadrupole. Thus, observation of the relative values of a number of zero-moment fields provides a measure of the relative strengths of the magnetic dipole and electric quadrupole interactions.

To find the exact theoretical peak positions and their dependence on Q, one solves the secular equation for the energy of the atom in an external field H directed, say, along the z axis. Omitting constant terms, the Hamiltonian is given by[4]

$$H'=\mu_0 H(g_J J_z+g_I I_z) +a\mathbf{I}\cdot\mathbf{J}+b2\mathbf{I}\cdot\mathbf{J}(2\mathbf{I}\cdot\mathbf{J}+1). \quad (2)$$

(The terms are, in order, the interaction between the external field and the electrons and nucleus, and the electron-nuclear interactions due to the nuclear magnetic dipole and electric quadrupole.) Since $I=9/2$ and $J=\frac{3}{2}$, the h. f. s. levels in zero field are characterized by F values of 3, 4, 5, 6; in a magnetic field each of these splits into $2F+1$ states, a total of 40 in all, giving a 40-rowed secular determinant.

However, in either the F, M_F representation or the M_I, M_J representation H' is diagonal in M_F, the z component of the total angular momentum; since at most four states have the same M_F the secular determinant factors into a series of smaller determinants having at most four rows, the solution of which leads to quartic equations in the energy. The F, M_F representation is the more convenient since in it the last two terms in Eq. (2) are diagonal with matrix elements given by $\frac{1}{2}aC+bC(C+1)$, from which Eq. (1) arises. The matrix elements of I_z and J_z are well known and are given by Condon and Shortley[5] or, in a notation which is the same as ours if $m=M_F$, in the appendix to reference 2.

It is convenient for calculation to express energies in terms of a; defining

$$x=g_J\mu_0 H/a, \quad (3)$$

we then have

$$H'=x\left(J_z+\frac{g_I}{g_J}I_z\right)+\mathbf{I}\cdot\mathbf{J} +\frac{b}{a}2(\mathbf{I}\cdot\mathbf{J})[2(\mathbf{I}\cdot\mathbf{J})+1]. \quad (4)$$

[2] Millman, Rabi and Zacharias, Phys. Rev. **53**, 384 (1938).
[3] Rabi, Kellogg and Zacharias, Phys. Rev. **46**, 157 (1934).
[4] H. B. G. Casimir, "On the Interaction between Atomic Nuclei and Electrons," Prize Essay published by Teyler's Tweede Genootschap (1936).
[5] E. U. Condon and G. H. Shortley, *Theory of Atomic Spectra* (Cambridge, 1935).

After calculating the numerical values of the matrix elements involving the angular momenta, the secular equation contains as parameters x, g_I/g_J, and b/a. Given any particular state, one desires the value of x, say x_0, at which $\partial E/\partial x=0$ (provided x_0 exists). x_0 will depend on the values of g_I/g_J and b/a; numerical calculation of this dependence is necessary—i.e., one assumes values for g_I/g_J and b/a, calculates $\partial E/\partial x$ for several values of x, and interpolates to find x_0. This process must be repeated to obtain x_0 for as many values of b/a and g_I/g_J as are necessary.

It is found that there is one value of the field at which the moment is zero for each of the following states, denoted by its (M_F, F) value: $(-4, 4)$, $(-4, 6)$, $(0, 4)$, and the twelve states for $M_F=-1, -2, -3$, $F=3, 4, 5, 6$. In addition there are two different zero-moment values of the field for the state $(1, 4)$.

The effect of g_I/g_J is very small, and, as will be discussed later, g_I/g_J is already known with an accuracy higher than can be obtained for it from the present experiment. In all the calculations we have therefore taken $g_I=-0.000774$, corresponding to $\mu_I=6.40$ nuclear magnetons,[2] and $g_J=4/3$ (see also later). For each peak, preliminary calculations of x_0 were made for $b/a=0$ and 0.006 and $dx_0/d(b/a)$ at $b/a=0$ was found. More exact locations of the zero moments were later calculated for $b/a=0.0056$, 0.0060, 0.0064. The results are shown in Fig. 1. No exact calculations were made on the low zero-moment field of the state $(1, 4)$ since it occurs at too low a field to be observed with the present apparatus.

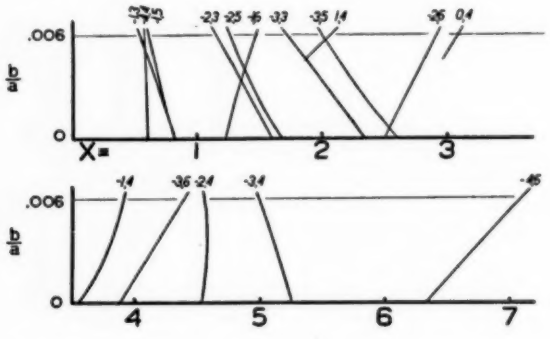


FIG. 1. Theoretical peak position as function of quadrupole interaction constant b/a. $x\sim 7.6\times 10^{-3}\,H$; see Eq. (2). Peaks denoted by (M_F, F) values.

Apparatus and Procedure

A schematic top view of the apparatus is shown in Fig. 2. Some features are similar to those discussed in reference 2, but to obtain a design with maximum resolution an analysis of the half-widths of zero-moment peaks was carried out. Unlike previous calculations these included the variation of the moment of the atoms in the beam across the width of the beam, an important factor at high resolution. We summarize briefly the analysis and its results.

Let H_0 be the value of the field at which the moment of the state under consideration is zero. To a close approximation, the absolute value of the deflecting field, H, varies linearly across the width of the beam and the magnetic moment at the field H varies linearly with $(H-H_0)$, i.e., with y, the distance from the zero-moment plane. (The locus of points for which $H=H_0$ is the "zero-moment plane," while the "beam plane" is determined by the centers of the source and collimating slits and of the detector. The two planes are parallel, and coincide at a zero-moment peak.) If $d^2E/dH^2>0$ for $H=H_0$, the zero-moment plane is a position of minimum energy and the atom is attracted to it by a force proportional to y. Such a state is called a "focusing" state. If $d^2E/dH^2<0$ ("nonfocusing" state), the atom is repelled from the zero-moment plane by a force proportional to y. From this it is easily shown that the atomic trajectories in the deflecting field are given by

$$y=y_0\cos\beta t+(\dot{y}_0/\beta)\sin\beta t \qquad \text{(focusing)}$$
$$y=y_0\cosh\beta t+(\dot{y}_0/\beta)\sinh\beta t. \qquad \text{(nonfocusing)}$$

y_0 and $\dot{y}_0$ are the values of y and the transverse velocity $\dot{y}$ at $t=0$;

$$\beta^2=\left|\frac{R^2H_0^2}{m}\left(\frac{d^2E}{dH^2}\right)_{H=H_0}\right|;$$

m is the mass of the atom and R is the ratio of field gradient to field.

As the value of the field at the beam plane is increased from less than to more than H_0, the deflection of a beam of atoms of the state in question decreases until it is zero for $H=H_0$ (and thus gives a peak in the intensity received at the detector) and then increases in the opposite direction. From the width of the deflected

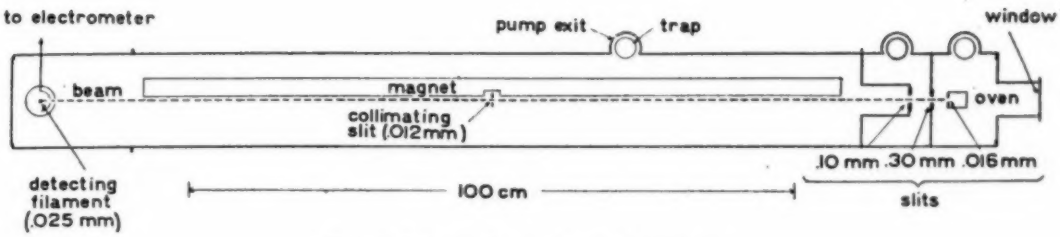


FIG. 2. Schematic top view of apparatus.

beam and the variation of its deflection with the value of H at the beam plane, both of which are easily calculated from the geometry of the apparatus and the above trajectories, the half-width of the peak (width at half the peak intensity), may be found for any given velocity of beam atoms.

The expression thus obtained is minimized with respect to the position of the collimating slit and the relative amounts of field-free space at each end of the magnet, with the result that the resolution is found to be best with a completely symmetrical arrangement such as that shown in Fig. 2. With this arrangement the half-width is given, for nonfocusing states, by the larger of the two following quantities:

$$(\Delta H/H)_c = RW_c\left[1-\frac{\lambda\gamma\cosh\gamma/2+2\sinh\gamma/2}{\lambda\gamma\cosh\gamma+(1+\lambda^2\gamma^2/4)\sinh\gamma}\right]^{-1};$$
$$(\Delta H/H)_s = \frac{RW_s}{2}\left[1-\frac{\lambda\gamma\cosh\gamma+(1+\lambda^2\gamma^2/4)\sinh\gamma}{\lambda\gamma\cosh\gamma/2+2\sinh\gamma/2}\right]^{-1}. \quad (5)$$

For focusing peaks one substitutes sin for sinh, cos for cosh, and changes the sign before the term $\lambda^2\gamma^2/4$. W_s and W_c are the source and collimating slit widths, respectively; λ is the ratio of total field-free region to length of the deflecting field, l; $\gamma=\beta l/v$ where v is the atomic velocity. The physical significance of γ is obvious; for $\gamma=\pi/2$, an atom of a focusing state which is traveling in the beam plane will, on entering the deflecting region, just reach the zero-moment plane when it leaves the deflecting region.

These equations are useful in finding the best practical combination of the remaining adjustable variables R, l, λ, W_s and W_c. The irregularities in the field itself set a limit to W_s and W_c of about 0.01 mm, below which any decrease in slit widths serves merely to decrease the intensity without helping the resolution. The final choice of R, l and λ is determined by such factors as the obtainable length for an accurately machined magnet, the length of beam convenient as to intensity of the beam and stability of the apparatus, etc. Roughly speaking, for good results one should have $\gamma\geq 1.2$ for focusing peaks and $\gamma\geq 1.5$ for nonfocusing peaks, with $\lambda\sim 0.3$. The half-width increases very rapidly as γ falls below l, unless λ is increased rapidly, which is usually impracticable.

A cross section of the deflecting magnet at right angles to the beam plane is shown in Fig. 3. The magnet consists of two copper tubes 0.144" *OD* and 0.066" *ID* set one above the other in the face of a Duralumin block 3.5×5×115 cm. (For such a magnet the quantity R previously referred to is approximately $1/a$ where a is the radius of the wires. All other quantities being constant, the resolution increases with R.) The tubes, insulated from one another and from the block by mica, carry the same current in opposite directions and are water cooled. After placing the wires in the block and before placing the block in the apparatus, the average deviation of the surface tangent to the wires from the mean tangent plane was about 0.007 mm. For the excitation of the field, currents up to 800 amperes (hereafter referred to as the "field current") were supplied by a six-volt bank of storage batteries of 3500 ampere-hours capacity. To measure the field current a Leeds & Northrup type K potentiometer and a 1500-ampere, 50-millivolt shunt were used. No direct measurements of the magnetic field were necessary; a much easier and more accurate calibration was made with an atomic beam of caesium, as will be described later. Because of this way of calibrating

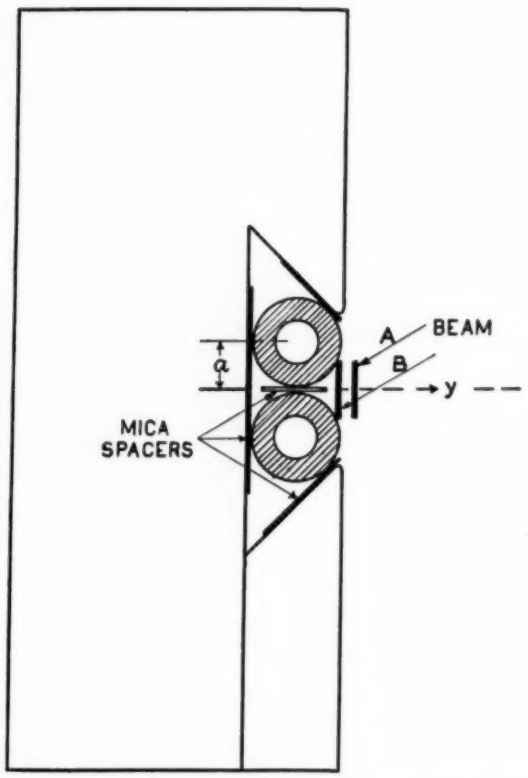


FIG. 3. Cross section of deflecting magnet. Field wires enlarged relative to Duralumin magnet block

the field, the calibration of the shunt does not affect the results.

The beam is defined by the oven and collimating slits, about 0.016 and 0.012 mm wide, respectively. The tungsten detecting filament is 0.025 mm in diameter. The height of the beam is limited to 2 mm by stops at each end. The oven and collimating slits are accurately vertical and may be moved under vacuum so that the beam may be placed parallel to the face of the magnet block and at any desired distance from it; the details of the adjustment are similar to those given in reference 2. One of the advantages of the symmetrical arrangement used here is that any error, within wide limits, in placing the beam plane parallel to the plane of the field wires has no effect whatsoever on the position or the half-widths of the observed peaks. (This does not apply, of course, to the verticality of the beam plane.) In the course of an experiment, the position of the beam at zero field was noted between each peak in order to avoid error caused by the walking of the beam; the position of the filament and hence the relative position of the beam could be observed to better than 0.0005 mm. Under good conditions the average shift of the beam during the observation of one or two peaks was 0.001 mm or less, and while corrections for this were calculated, they were in most cases negligible.

Data were taken over a number of runs on three different line-ups, the first two with the beam in position A as shown in Fig. 3, about 0.6 mm outside the plane defined by the surface of the field wires and the third 0.1 mm inside, at position B.

EXPERIMENTAL DATA

Figure 4 shows the complete course of the beam intensity as a function of field current. The large background is due to $^2P_{1/2}$ atoms, largely in the state $M_F=0$, which have smaller over-all moments than the $^2P_{3/2}$ and therefore are not swept out of the field as readily. It is known from the work of Millman, Rabi and Zacharias on the

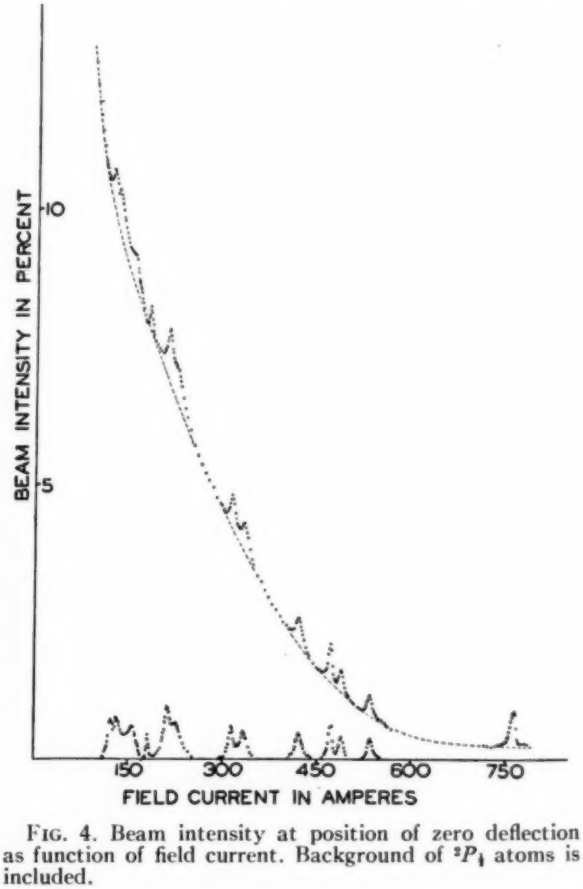


FIG. 4. Beam intensity at position of zero deflection as function of field current. Background of $^2P_{\frac{1}{2}}$ atoms is included.

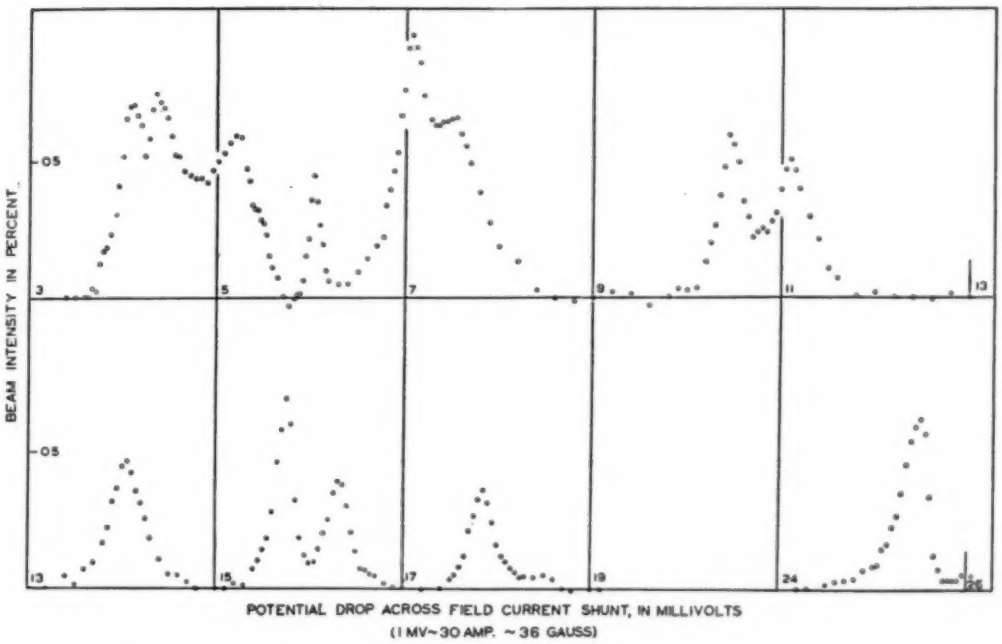


FIG. 5. Beam intensity, at position of zero deflection, of $^2P_{3/2}$ atoms alone.

ground state that there are no $^2P_{3/2}$ peaks at higher fields than those shown. In locating the $^2P_{3/2}$ peaks the $^2P_{1/2}$ background is subtracted by running a smooth curve through the points adjacent to the peaks. The peaks thus obtained are shown in Fig. 4 and in more detail in Fig. 5.

The unequal intensity of the peaks is not caused by unequal population of the corresponding states. At a peak arising from a focusing state the atoms travel down the middle of a long potential valley to the detector and are focused in to the detector to give an intense peak. For a nonfocusing peak the atoms are trying to travel a straight and narrow path along the top of a potential ridge and much fewer reach the detector. The area under a peak is not a measure of population either, as the half-widths are not necessarily greater for the nonfocusing than for the focusing peaks. This may be seen with reference to Eqs. (5).

An asymmetry may also be observed for several peaks, notably the last one. This occurs because these curves were taken on line-up *B* as shown in Fig. 3 so that the wires constitute, in effect, an asymmetrical height stop; because of this, focusing peaks have their high field intensity, the nonfocusing peaks their low field intensity reduced. This asymmetry has been taken into account in locating the peaks. It does not appear in the curves taken with the line-ups for which the beam is clear of the wires.

The directly observed quantity which gives the position of a peak is the potential drop, in millivolts, across the field current shunt, hence there were three observed positions of a given peak corresponding to the three different selections of position of the beam with respect to the field wires. In order to compare the peaks of the first two line-ups with those of the third line-up, for which the resolution was best, the ratios of the positions (in millivolts) of each peak on line-up 1 to the positions of the same peaks on line-up 3 were calculated and a weighted average over all peaks gave a conversion factor changing millivolts on line-up 1 to millivolts on line-up 3. A similar procedure was followed for line-up 2. The resulting comparison showed no discrepancies, beyond the limits of error which had previously been assigned for the location of the peak positions in millivolts, between the results obtained on the three line-ups, except in one case. Here there had been a systematic error in locating two peaks which were poorly resolved on both the first two line-ups.

From the three positions of each peak and the corresponding limits of error a final position and

limit of error was set. The numbers which enter into the final results of the experiment were based on an average of 16 observations for each peak, distributed over 10 runs. The results are shown in Table I.

Results

The relative positions at which the peaks occur depend only on the ratio b/a and the nuclear and electronic total angular momenta, I and J. The absolute value of the magnetic field at each peak is proportional to a/g_J, i.e., to the relative strength of the electron-nuclear and electron-external field magnetic interactions. From the numbers upon which Fig. 1 is based, the dependence on b/a of a number of ratios of peak positions was calculated and from the experimentally observed values of these ratios the corresponding values of b/a were determined. Of the resulting values of b/a, those having limits of error of less than three percent are shown in Fig. 5. The weighted average of these gives $b/a=(6.432\pm0.04)\times10^{-3}$. Here, as in all further results, the plus or minus represents a reasonable limit of error. It will be noted that the value of b/a is entirely independent of any measurement of absolute value of magnetic field.

The hyperfine structure separation

Knowing b/a, we may now determine from the calculations on which Fig. 1 is based the value of x at which any given peak occurs. A measure of the field at this peak then gives a/g_J directly, by use of Eq. (3). The most accurate way of calibrating the field is to run another element whose peaks occur at known fields. In this case caesium was used. The $\Delta\nu$ of the $^2S_{1/2}$ ground state of caesium is known to be 0.3067 ± 0.0004 cm^{-1} spectroscopically;[6] Millman and Fox,[7] using atomic beam methods, found 0.307 ± 0.003 cm^{-1} for $\Delta\nu$ and 7/2 for the spin. Given the $\Delta\nu$, for which we take the spectroscopic value, and the spin, the field at any peak of caesium is then known. It was not found feasible to obtain caesium and indium beams from the same source oven, so the caesium run was bracketed between two indium runs. The difference in the position of the $(-3, 4)$ indium peak (the most accurately reproducible of the indium peaks) on the two runs was at most 0.1 percent and the maximum spread of locations of this peak on three indium runs preceding the caesium run over a period of three weeks was 0.06 percent. Experimentally, the ratio of the $(-3, 4)$ peak to the first caesium peak was found to be 0.7913 ± 0.0022. The $(-3, 4)$ peak comes at $x=4.9839\pm0.0025$ for the value of b/a found above. Making use of these two numbers and the $\Delta\nu$ and spin of caesium we obtain, from Eq. (3) and the analogous equation for caesium, $a/g_J=(6.087\pm0.028)\times10^{-3}$ cm^{-1}. If one assumes for g_J the value for a pure $^2P_{3/2}$ state, 4/3, this gives $a=(8.116\pm0.037)\times10^{-3}$ cm^{-1}. By way of comparison, Schüler and Schmidt,[8] who also investigated the h. f. s. of this state, give $a=7.97\times10^{-3}$ cm^{-1}, obtained from the two largest separations of the h. f. s. pattern. Their separations are stated to 0.001 cm^{-1}, and an error of this amount in one of the measured separations corresponds to an error of 0.15×10^{-3} cm^{-1} in a. The two values are not in contradiction in view of their respective limits, but because of the possible influence of perturbations on g_J (see later), we increase the limit of error and take $a=(8.10\pm0.10)\times10^{-3}$ cm^{-1}. If there were no quadrupole moment the h. f. s. separation would then be given by $\Delta\nu=15a=(0.1215\pm0.0014)$ cm^{-1}. Because of the presence of the quadrupole, which shifts the two extreme levels by different amounts, the actual separation of these levels is (0.1184 ± 0.0015) cm^{-1}, making use of the value of $b=(0.0521\pm0.0009)\times10^{-3}$ cm^{-1} deduced from the above values of b/a and a. Schüler and Schmidt found $b=0.048\times10^{-3}$ cm^{-1}, a value which is altered by ten percent by a

TABLE I. *Final determinations of peak positions.*

Peak (M_F, F)	Peak position mv on line-up 3	Peak (M_F, F)	Peak Position mv on line-up 3
−2,3	4.068±0.035	0,4	11.122±0.040
−2,5	4.410±0.04		
−1,6	5.215±0.06	−1,4	14.020±0.025
−3,3	6.040±0.015	−3,6	15.760±0.020
		−2,4	16.300±0.025
−3,5	7.080±0.020	−3,4	17.836±0.023
1,4	7.500±0.035	−4,6	25.44 ±0.10
−2,6	10.463±0.026		

[6] L. P. Granath and R. K. Stranathan, Phys. Rev. **48**, 725 (1935).
[7] S. Millman and M. Fox, Phys. Rev. **50**, 220 (1936).
[8] H. Schüler and T. Schmidt, Zeits. f. Physik **104**, 468 (1937).

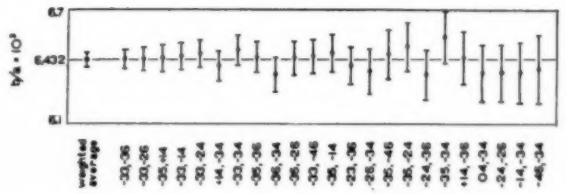


FIG. 6. Values of b/a calculated from observed peak ratios. The two peaks contributing to a given value of b/a are specified by their (M_F, F) values.

change of 0.001 cm^{-1} in one of their measured separations.

Nuclear spin and nuclear g factor

The fact that an unambiguous identification of observed and theoretical peaks may be made serves to determine the correctness of the previously measured value 9/2 for the nuclear spin I, upon which the calculations were based.

Strictly speaking, the ratio of the nuclear and electronic g factors, which occurs in Eq. (4), affects slightly the relative positions of the peaks. It was because of this fact that Millman, Rabi and Zacharias[2] were able to find the nuclear g factor of indium directly rather than through calculations based upon h. f. s. separations. To a very close approximation, the shift of a peak by the term in g_I may be obtained from the following considerations. If g_I (and hence the nuclear moment) is zero then the zero-moment value of the field is that for which **J** is, on the average, perpendicular to **H**—i.e., for which $\bar{M}_J=0$. Since M_F is a constant of the motion and $M_F=M_I+M_J$, it follows that $\bar{M}_I=M_F$. If now g_I is not zero, the moment of the atom at this particular field is $-g_I\bar{M}_I\mu_0=-g_IM_F\mu_0$, which is of the order of 0.001–0.002 Bohr magneton for the case in question. If one knows $d\mu/dH$, which is easily calculated, then ΔH_0, the shift in the zero-moment value of the field due to g_I, is given by $\Delta H_0=g_IM_F/(d\mu/dH)$. The electronic moment is 2 Bohr magnetons for the $^2P_{3/2}$ state as compared with $\frac{1}{3}$ of a Bohr magneton for the $^2P_{1/2}$, and the effect of this factor in increasing $d\mu/dH$ is to make the percent shift of the peaks correspondingly smaller for the $^2P_{3/2}$ state. Since the g_I of Millman, Rabi and Zacharias is therefore more accurate than one which could be obtained from the present experiment, we have used their value in calculating the theoretical peak positions shown in Fig. 1. If we set $g_I=0$ in the theoretical calculations, the resulting changes in the values of b/a as deduced from the experimental data are in no case changed by more than the limits of error shown in Fig. 6. The consistency of the values of b/a, however, is made noticeably worse, indicating that the assumed sign and order of magnitude of g_I are correct.

The quadrupole moment

For this particular state, the relation between Q and b is[4, 5]

$$Q=b\frac{I(2I-1)Z_i^3HC}{\delta R'}\frac{60\mu_0^2}{e^2},$$

Here δ is the doublet separation, 2212.6 cm^{-1}, which enters since $\overline{r^{-3}}$ is proportional to it. H, R', and C are relativistic correction factors which enter because of the use of δ to give $\overline{r^{-3}}$. Calculations of these quantities have been made by Casimir,[4] from whose tables we obtain $H=1.050$ and $R'=1.088$. C is a correction for the fact that in calculating H and R' the difference in the normalization integrals of the $^2P_{1/2}$ and $^2P_{3/2}$ states, caused by the difference in the energy of these two states, was neglected. Casimir gives an analytical expression for this effect from which we obtain $C=1.025$. Z_i is the "effective" nuclear charge. On the basis of the relativistic wave functions from which H, R' and C are calculated, Casimir obtains an expression for δ which depends on Z_i. From the doublet separations and term values, E_n, of the series $5s^2np$, a value of Z_i may be obtained for each n. Averaging the values obtained from $n=6$ to 10 inclusive we find $Z_i=44.8=Z-4.2$, with an average deviation of 1.0. (The value for $n=5$ was omitted since its calculation involved a rather crude estimation of dE_n/dn, which is the form in which E_n enters the formula for δ.) Finally, in the calculation of μ_0/e, in which μ_0 and e are the Bohr magneton and the electronic charge, we have taken $h/e=1.3759\times10^{-7}$ e. s. u. and $e/m=1.7591\times10^7$ e. m. u. b and δ are expressed in cm^{-1}.

Substitution of these numerical values gives $Q=0.837\times10^{-24}$ cm^2; we take

$$Q=0.84\times10^{-24}\ cm^2.$$

Some attention must be paid to perturbations of the $5s^25p\ ^2P_{3/2}$ state in question by electrostatic

coupling with states of the same total angular momentum in higher configurations, notably the three states with $J=\frac{3}{2}$ in the $5s5p6s$ configuration. As was shown by Fermi and Segrè[9] in the closely analogous case of thallium, such coupling is of the right order of magnitude to explain the anomalously small $\Delta\nu$ of this state, which is about half what it should be on the basis of the known value for the magnetic moment. While the levels of the $5s5p6s$ configuration have not been observed spectroscopically, the calculations of Fermi and Segrè indicate that it lies roughly 50,000 cm^{-1} above the ground level, with a matrix element of the electrostatic energy of the order of 5000 cm^{-1} coupling the two configurations. This means that properties of the states in the $5s5p6s$ configuration mix with those of the lower level in the ratio of about 1 : 100. Since the upper configuration has two unbalanced s electrons the $\Delta\nu$'s of the various multiplet levels will be very large. If we represent the total effect by a single $\Delta\nu$ with a reasonable value of 12 cm^{-1} and an inverted pattern, the $\Delta\nu$ of the lower state is cut in half as is observed. The situation is quite different as regards the quadrupole moment, however; s electrons, having spherically symmetrical wave functions, do not exhibit the quadrupole effect in h. f. s. The quadrupole contributions to the energy levels of both configurations arise from the $5p$ electron and hence will be of the same order of magnitude so that perturbations in the value of Q observed for the lower state will have an order of magnitude of at most one percent. As concerns g_J, previously referred to, the g sum rule shows that the average g_J for the states $J=\frac{3}{2}$ in the upper level is $22/15=1.47$ as compared with $4/3$ for the state in question, so that g_J might possibly be increased by 0.3 percent.

Discussion

Schüler and Schmidt, whose work on the same state has already been mentioned, obtained for 0.8 ± 0.2. The large probable error arises from the fact, previously noted, that a 0.001-cm^{-1} error in one of their h. f. s. level separations causes a ten-percent error in b, and hence in Q. Bacher and Tomboulian,[10] working with the $5s6p\ {}^1P_1$ state of In II, found $Q=0.82$. The agreement with our value of 0.84 is very pleasing; but in view of the fact that the method of calculating Q from h. f. s. is very similar to that used for nuclear moments, which usually show a wide spread, the agreement may be in part fortuitous. Much more significant is the internal consistency of our results, as indicated by the 0.6 percent limit of error on b/a. Because there are thirteen observables in this experiment as compared with the much smaller number of h. f. s. energy differences observable in a spectroscopic investigation of a single state, this consistency provides a much more complete check on the sufficiency of the assumed interaction between nucleus and electrons.

In this connection, Tolansky[11] has recently reported that he finds it necessary, for several states of iodine, to add to Eq. (1) a term $c(C^3+4C^2+4/5C)$ representing a cosine^3 interaction between nucleus and electrons. The constant c is presumably proportional to a nuclear magnetic octopole moment, and c/a is of the order of magnitude of 2×10^{-4}. We have calculated the effect which such a term would have on our peaks; we find that if $c/a=5\times10^{-6}$, and if the term in c/a is neglected in calculating the theoretical dependence of peak position on b/a, then some of the more accurately determined experimental values of $b/a\times10^3$ shown in Fig. 6 should be inconsistent by $1\frac{1}{2}$ percent, which is not so for the values in question. Furthermore, such inconsistencies as appear in Fig. 6 are not bettered by any choice of c/a. Hence 5×10^{-6} seems a good upper limit to c/a in this state.

In conclusion, the author desires to acknowledge his indebtedness to Professor I. I. Rabi, who suggested this application of molecular beam methods to the measurement of quadrupole moments and whose advice was very helpful. He also wishes to express his deepest gratitude to Mr. Nicholas A. Renzetti for his continuously invaluable assistance during the course of the research. The other members of the molecular beam laboratory have been of great aid in contributing the benefits of their previous experience.

[9] E. Fermi and E. Segrè, Zeits. f. Physik **82**, 729 (1933).
[10] R. F. Bacher and D. H. Tomboulian, Phys. Rev. **52**, 836 (1937).
[11] S. Tolansky, Proc. Roy. Soc. **A170**, 205 (1939).

Stark-Effect Broadening of Hydrogen Lines

II. Observable Profiles

LYMAN SPITZER, JR.*
Harvard College Observatory, Cambridge, Massachusetts
(Received May 4, 1939)

The hydrogen-line profiles previously derived for the case of single encounters are integrated over all encounters to yield the observable profiles. The encounters are assumed to be uniformly distributed over different distances of closest approach, and the ions are assumed to have a Maxwellian velocity distribution. Approximation functions allow the integration to be carried through with rigor in the two limiting cases of very fast and very slow encounters, and with a probable error not exceeding ten percent in the more general case. The resultant profiles illustrate quantitatively Margenau's effect—the transition from impact to statistical broadening with increasing line width. The effect of fast electron encounters is shown to be negligible, in general, in comparison with the effect produced by an equal density of slowly moving ions. An evaluation is made of the error produced by the neglect of the inhomogeneity of the ionic field. The final formulas have a probable error of less than 12 percent for most cases of astrophysical interest.

THE perturbations which a passing ion produces on a hydrogen atom have recently been investigated by the author.[1] The profile of a line emitted by an atom so perturbed has been derived on the assumptions that the field of the ion is homogeneous and that the motion of the ion is uniform and rectilinear. The results of I may be conveniently summarized as follows. Let $J(x)$ be the intensity of emitted radiation as a function of x, where x equals $2\pi(\nu-\nu_0)$, and is simply the distance in angular frequency units from the line center ν_0. Let Γ be the radiative damping constant for the transition in question, and let $\pi\delta$ be the total phase shift produced by the encounter. When x is much greater than Γ, formula (20-I) is valid, provided that $f_\sigma(\xi)$ is substituted for $f_\delta(\xi)$ in accordance with Sections 2 and 3 of I; thus we have

$$J(x)=\frac{\Gamma}{2\pi x^2}\left\{1+\frac{\Omega(\delta,\xi)}{\Gamma}\delta^2 f_\sigma^2(\xi)\right\}, \qquad (1)$$

where

$$\delta=q/\hbar vR, \qquad (2)$$

$$\sigma^2=1+\delta^2, \quad \xi=xR/v, \qquad (3)$$

and $\Omega(\delta,\xi)$ is the number of collisions per second for a particular δ and ξ; v is the velocity of the ion relative to the hydrogen atom, and R is the distance of closest approach between the ion and the atom. The function $f_\sigma(\xi)$ is defined by (21-I), and is represented graphically by Fig. 1 in I. It is evident from this figure that if σ is much greater than unity, $f_\sigma(\xi)$ is given by the statistical formula (30-I) derived for infinitely slow encounters. When δ is less than unity, σ is substantially unity and $f_\sigma(\xi)$ approaches the known limiting curve $f_1(\xi)$ given in (28a-I).

Strictly speaking, we have proved (1) only for the two limiting cases of δ large and δ small. The discussion of the lowest Lyman line given in Section 2 of I validates the functional form of (1), however, and it seems reasonable to expect that this formula is substantially correct for intermediate values of δ in other cases as well.

The above formulation assumes that all encounters are characterized by the same values of δ and ξ. In the more general case we must integrate (1) over all possible collisions to find the profile of a line emitted by a hydrogen atom in an assembly of ions. The actually observed line contour will be the same as this theoretical "observable profile" only when the emission takes place from a thin film and self-absorption is therefore negligible. This integration is carried out in Section 1 below. In Section 2 the results are qualitatively discussed and their range of validity is investigated.

1

The integration of (1) over all possible values of δ and ξ may be regarded as an integration over R and v. If we let $\Omega_m'(R,v)dRdv$ be the

* National Research Fellow.
[1] L. Spitzer, Jr., Phys. Rev. **55**, 699 (1939), hereafter referred to as I.

number of times per second that a single hydrogen atom encounters an ion of mass M_m, of charge $Z_m e$, with a relative velocity between v and $v+dv$, and with a distance of closest approach between R and $R+dR$, then (1) becomes

$$J(x)=\frac{\Gamma}{2\pi x^2}\left\{1+\frac{1}{\Gamma}\sum_m\int_0^\infty dv\int_0^s \times\Omega_m'(R,v)\delta^2 f_\sigma^2(\xi)dR\right\}, \quad (4)$$

summed over all types of ion. The upper limit of integration over R in (4) is the average distance of the nearest ion of any type and is defined by

$$4\pi s^3/3=1/N, \quad (5)$$

where N is the total number of ions per cm³. This is perhaps an arbitrary cut-off but represents roughly the fact that encounters such that R is greater than s will in general be completely screened. Fortunately the relevant results depend only logarithmically on this upper limit of integration s. In this way a consideration of successive, isolated, single encounters can be made to provide an adequate approximation to the actual situation.

The lower limit of integration over R has been set equal to zero in (4), although the theory in I breaks down for sufficiently small R, owing to the inhomogeneity of the ionic field, exchange forces, and other phenomena. It seems a better procedure, however, to include an admittedly inaccurate value for the contribution from these close encounters rather than to neglect them entirely. The error introduced from this source is discussed more fully in Section 2 below.

To obtain an explicit formula for $\Omega_m'(R,v)$ in (4) one may assume a Maxwellian velocity distribution for the relative motions of ions and hydrogen atoms. As before, the effect of the perturbation on the ionic motion may be neglected, and the simple geometrical cross section may be used in determining the frequency of collisions. With these two assumptions, $\Omega_m'(R,v)$ becomes

$$\Omega_m'(R,v)=2\pi RvN_m\cdot 4l_m^{\frac{1}{2}}\pi^{-\frac{1}{2}}v^2\exp -l_m v^2, \quad (6)$$

where

$$l_m=\frac{1}{2kT}\frac{M_m m_{\mathrm{H}}}{M_m+m_{\mathrm{H}}}; \quad (7)$$

m_{H} is the mass of the hydrogen atom, and N_m is the number of mth ions per cm³.

The substitution of (6) into (4) and the use of (2) for δ yields an explicit integral formula for $J(x)$. This may be simplified by substituting u for $l_m v^2$ and by reintroducing ξ for Rx/v. This gives

$$J(x)=\frac{\Gamma}{2\pi x^2}\left\{1+\frac{4\pi^{\frac{1}{2}}}{\Gamma x^{\frac{1}{2}}}\sum_m\frac{N_m q_m^{\frac{1}{2}}}{\hbar^{\frac{1}{2}}}\gamma_1^{\frac{1}{2}}H_m(x)\right\}, \quad (8)$$

where

$$H_m(x)=\int_0^\infty e^{-u}du\int_0^{\gamma_1/\gamma_2 u^{\frac{1}{2}}} f_\sigma^2(\xi)d\xi/\xi, \quad (9)$$

and from (2) and (3),

$$\sigma^2=1+\gamma_1^2/u^2\xi^2, \quad (10)$$

while by definition we have

$$\gamma_1=q_m l_m x/\hbar, \quad \gamma_2=q_m l_m^{\frac{1}{2}}/\hbar s. \quad (11)$$

The quantity q_m/h is the Stark-effect shift in frequency units which a stationary charge $Z_m e$, placed at unit distance from the hydrogen atom, will produce in the particular component under consideration.

The dimensionless quantities γ_1 and γ_2, from which a subscript m is omitted, have each an important physical significance. For an ion with the most probable relative velocity $l_m^{-\frac{1}{2}}$, γ_1 equals the product $\delta^2 x/x_{\max}$, where

$$x_{\max}=q_m/\hbar R^2; \quad (12)$$

$x_{\max}$ is simply the maximum value of x for a given R in the limiting statistical case. Hence if γ_1 is small, δ for most collisions will be greater than unity only if $x/x_{\max}$ is very small. For sufficiently small values of x, $f_\sigma(\xi)$ equals $2\sin(\pi\sigma/2)/\sigma$, as is evident from (21-I), while for larger x δ will be small, if γ_1 is still small, and $f_\sigma(\xi)$ becomes substantially $f_1(\xi)$. If, on the other hand, γ_1 is large, δ will be so great that $f_\sigma^2(\xi)$ is given by the limiting form (30-I) until $x/x_{\max}$ is so large that $f_\sigma(\xi)$ vanishes in any case. Thus for large γ_1 we may expect the usual statistical formula to apply.

It follows from (2) that γ_2 in (11) is the value of δ for an encounter with an ion of velocity $l_m^{-\frac{1}{2}}$ and closest distance of approach s. If this is greater than unity all relevant encounters produce phase shifts greater than unity, since from (2) δ increases as R decreases. We have seen

that when σ is much greater than unity, $f_\sigma(\xi)$ approaches the limiting statistical curve $f_\infty(\xi)$. The full statistical theory may therefore be used when γ_2 is large, and the analysis of Holtsmark[2] and Margenau[3] may be applied.

The ratio γ_1/γ_2^2 equals $xs^2\hbar/q$ and is simply the ratio x/x_{max} when R is equal to s. If all ions have the same charge Ze, this is practically equal to Holtsmark's[2] ratio β, or F/F_n, where F is the electrical intensity whose probability distribution at a point is under consideration, and F_n is a certain normal or average intensity defined by

$$F_n=\{8\pi(2\pi)^{\frac{1}{2}}/15\}^{\frac{4}{7}}ZeN^{\frac{2}{3}}. \tag{13a}$$

If we substitute for N from (5), this gives

$$F_n=\{2(2\pi)^{\frac{1}{2}}/5\}^{\frac{4}{7}}Ze/s^2=1.0018Ze/s^2. \tag{13b}$$

Thus F_n is essentially the field produced by a charge Ze at a distance s; $q/s^2\hbar$ will be the value of x produced by such a stationary charge—see (12)—and hence it follows that β equals x/x_{max} when R is equal to s. In other words, β is simply the ratio of the line shift x to the shift produced by a stationary ion at the average interionic distance s. This quantity β is of importance below in a comparison of the results of the statistical analysis with those of the more general treatment.

The evaluation of the double integral in (9) to obtain $H_m(x)$ is the central problem of the present analysis and gives directly the observable profile $J(x)$. The indicated integration can apparently not be carried out exactly in terms of known functions, particularly since $f_\sigma(\xi)$ is determined in I only for integral values of σ, and for other values must be found by interpolation. An approximate treatment may be adopted. It is evident from the graph of $\delta f_\delta^2(\xi)$ against x/x_{max} in I that while this function changes continuously from δ equal to one to δ equal to infinity, the difference between the curve for δ equal to four and that for δ equal to infinity is not very great. On the other hand, for x/x_{max} small, $f_\sigma(\xi)$ is given more accurately by $f_\sigma(0)$. We may approximate to $f_\sigma(\xi)$, then, in the following manner: when x/x_{max} is less than $\frac{1}{16}$ we assume

$$f_\sigma(\xi)=f_\sigma(0); \tag{14}$$

when x/x_{max} is greater than $\frac{1}{16}$, however, we set

$$f_\sigma(\xi)=\begin{cases} f_1(\xi) & 1 <\sigma<1.5, \\ f_2(\xi) & 1.5<\sigma<3, \\ f_4(\xi) & 3 <\sigma<5, \\ f_\infty(\xi) & 5 <\sigma. \end{cases} \tag{15}$$

The choice of $\frac{1}{16}$ as the critical value of x/x_{max} is somewhat arbitrary. As was shown above, the value of δ at this point will be $4\gamma_1^{\frac{1}{2}}$ for an ion with the relative velocity $l_m^{-\frac{1}{2}}$. If this value of δ is greater than three, the contribution of (14) to (9) will be not greater than four percent of the total, and thus quite negligible. For small γ_1, on the other hand, the transition at x/x_{max} equal to $\frac{1}{16}$ will take place directly between $f_\sigma^2(0)$ and $f_1^2(\xi)$, and both will be very nearly equal to four, as may be seen from the form of these functions as given in (20) below. The transition in this case is therefore continuous and independent of the particular transitional value of x/x_{max} chosen.

When β is less than $\frac{1}{16}$, x/x_{max} is always less than $\frac{1}{16}$ and $f_\sigma(0)$ replaces $f_\sigma(\xi)$ for all u and ξ. In this case the integrals in (9) are independent of x. We shall assume throughout that β is greater than $\frac{1}{16}$ and deduce the results for this other case by taking the limit as β decreases to $\frac{1}{16}$. It is apparently necessary to assume also that γ_2 is considerably less than unity, say less than 0.2. When γ_2 is very small, simultaneous encounters are not very important, and deviations from the statistical assumptions may be considered. When γ_2 is only slightly less than unity, and γ_1 is also less than unity, both effects are important and the present theory breaks down.

With these simplifications the integration in (9) becomes straightforward. We may define

$$L(x, u)=\int_0^\infty f_\sigma^2(\xi)d\xi/\xi; \tag{16}$$

then we find from (9), omitting the subscript m when there is no ambiguity

$$H(x)=H^1(x)-H^2(x), \tag{17}$$

$$H^1(x)=\int_0^\infty e^{-u}L(x, u)du, \tag{18}$$

$$H^2(x)=\int_0^\infty e^{-u}du\int_{\gamma_1/\gamma_2u^{\frac{1}{2}}}^\infty f_\sigma^2(\xi)d\xi/\xi. \tag{19}$$

[2] J. Holtsmark, Ann. d. Physik **58**, 576 (1919).
[3] H. Margenau, Phys. Rev. **43**, 129 (1933); H. Margenau and W. W. Watson, Rev. Mod. Phys. **8**, 22 (1936).

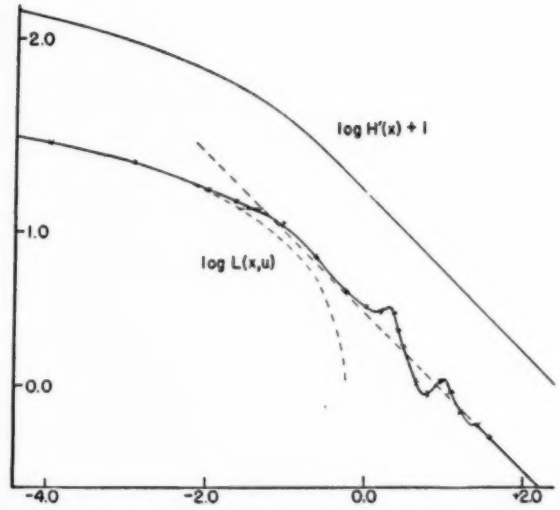


FIG. 1. The lower solid-line curve shows log $L(x, u)$ as a function of log γ_1/u. The two asymptotic forms are shown by the dashed-line curves. The upper curve gives log $H^1(x)$ $+1.00$ as a function of log γ_1. Both ordinates and abscissae are common logarithms.

The functions in (14), (15) and (16) may be obtained from (27-I), (28a-I) and (30-I); we find

$$f_\sigma^2(0)d\xi/\xi=\frac{4\sin^2\pi\sigma/2}{\sigma(\sigma^2-1)}d\sigma, \tag{20a}$$

$$f_1^2(\xi)d\xi/\xi=4\xi\{K_1^2(\xi)+K_0^2(\xi)\}d\xi, \tag{20b}$$

$$f_2^2(\xi)d\xi/\xi=2\pi^2\xi e^{-2\xi}d\xi, \tag{20c}$$

$$f_4^2(\xi)d\xi/\xi=2\pi^2\xi(1-\xi)^2e^{-2\xi}d\xi, \tag{20d}$$

$$f_\infty^2(\xi)d\xi/\xi=\begin{cases}\dfrac{\pi u^{\frac{1}{2}}}{2\gamma_1^{\frac{1}{2}}}\dfrac{d(x/x_{\max})}{(1-x/x_{\max})^{\frac{1}{2}}} & x/x_{\max}<1,\\ 0 & x/x_{\max}>1.\end{cases} \tag{20e}$$

These functions have been averaged over positive and negative values of ξ. The resultant profiles must accordingly be summed over all states, independently of the sign of q_m.

The limits implied in (14) and (15) may be expressed in terms of σ and ξ if we use (10) and the formula

$$x/x_{\max}=u\xi^2/\gamma_1, \tag{21}$$

which may be derived from (3), (11) and (12). If the denominator of (20a) is resolved into partial fractions, the integral can then be evaluated in terms of logarithms and the cosine integral $Ci(x)$, which is tabulated in Jahnke and Emde.[4] The second function integrates directly to $4\xi K_1(\xi)K_0(\xi)$, as may be seen by differentiation and application of the recurrence formulas for Bessel functions. The other integrations are elementary.

The resultant curve for $L(x, u)$ is shown by the lower solid curve in Fig. 1. The two dashed lines are the asymptotic curves,

$$L(x, u)=\pi u^{\frac{1}{2}}/\gamma_1^{\frac{1}{2}} \tag{22a}$$

for large u, which follows directly from (20e), and the series expansion for small u,

$$L(x, u)=4\ln u/\gamma_1-0.69, \tag{22b}$$

where $\ln u/\gamma_1$ denotes the logarithm of u/γ_1 to the base e.

Consideration of the defining integral (16) indicates that the oscillations of $L(x, u)$ shown in Fig. 1 are probably not real, since $f_\sigma^2(\xi)/\xi$ is a function which oscillates from zero through positive values with an amplitude that increases slowly with ξ until finally σ approaches unity and the function decreases monotonically. An increase in γ_1/u will shift any particular zero or maximum of the function to greater ξ but should not produce any rapid oscillations in the integral. The oscillations shown in Fig. 1 can in any case be traced to defects in the approximations (14) and (15) and are produced by the joining of two of the functions used in such a way that both functions have maxima, or both have minima, in their respective ranges.

If, then, we draw a smooth curve through the oscillations of $L(x, u)$ in Fig. 1, we see that the resultant curve follows closely either one or the other of the two asymptotic curves. The two dashed lines, connected by the dotted horizontal line, form an approximation curve whose vertical distance from the curve for $L(x, u)$ never exceeds 0.05 on the logarithmic scale, and which therefore represents $L(x, u)$ with an accuracy of about 12 percent for the short transition region and is of course correct at the two limits.

The connecting horizontal line is defined by

$$L(x, u)=13.80, \quad 0.0520>\gamma_1/u>0.0266. \tag{23}$$

If this and the two asymptotic curves are

[4] E. Jahnke and F. Emde, *Tables of Functions* (B. G. Teubner, 1933), p. 78.

integrated over their respective ranges in (18), $H^1(x)$ becomes

$$H^1(x)=\pi^{\frac{1}{2}}\,\mathrm{erf}\,(4.39\gamma_1^{\frac{1}{2}})/2\gamma_1^{\frac{1}{2}}+4Ei(37.4\gamma_1), \quad (24)$$

where
$$\mathrm{erf}\,(x)=\frac{2}{\pi^{\frac{1}{2}}}\int_0^x \exp(-t^2)dt, \quad (25a)$$

$$Ei(x)=\int_x^\infty \exp(-t)dt/t. \quad (25b)$$

These functions are also tabulated in Jahnke-Emde. When γ_1 is greater than 0.10 or less than 0.03, (24) becomes, respectively,

$$H^1(x)=\pi^{\frac{1}{2}}/2\gamma_1^{\frac{1}{2}}, \quad (26a)$$

$$H^1(x)=4\ln 1/\gamma_1-3.02+61\gamma_1, \quad (26b)$$

with an accuracy better than five percent. This function is shown by the upper curve of Fig. 1 and its values are given in Table I.

To determine $H(x)$ and hence $J(x)$ we must also calculate the correction term $H^2(x)$ given by (19). At the lower limit of integration over ξ in (19), $x/x_{\max}$ and σ are equal to β and to $(1+\gamma_2^2/u)^{\frac{1}{2}}$, respectively, as is evident from (21) and (10). From (14) and (15) it follows that if we take β greater than $\frac{1}{16}$ and γ_2 less than 0.2—the restrictions assumed above—$f_\sigma(\xi)$ will be replaced by $f_1(\xi)$ if u is greater than 0.032. The contribution to (19) from values of u less than 0.032 will be quite negligible, and we may therefore replace $f_\sigma^2(\xi)$ by $f_1^2(\xi)$ for all values of u and ξ in (19).

The indefinite integral over ξ in (19) thus equals $4\xi K_1(\xi)K_0(\xi)$, as we have already seen. This is not a convenient function to integrate over u. If we make the approximation

$$\xi K_1(\xi)K_0(\xi)\approx Ei(\tfrac{1}{2}e^{\frac{1}{2}\pi}\xi)+\tfrac{1}{2}\pi e^{-2\xi}, \quad (27)$$

where $Ei(x)$ is given in (25b), the functions are equal for large ξ and for small ξ, and differ nowhere by more than six percent. With this approximation, then, we have for (19), integrating the first term by parts,

$$H^2(x)=2\{F(1.447\gamma_3)-\pi\gamma_3F'(\gamma_3)\}, \quad (28)$$

where by definition

$$\gamma_3=4\gamma_1^2/\gamma_2^2, \quad (29)$$

$$F(z)=\int_0^\infty \exp(-zu-u^{-1})du/u, \quad (30)$$

$$-F'(z)=\int_0^\infty \exp(-zu-u^{-1})du. \quad (31)$$

Values of $F(\gamma_3)$, $\gamma_3F'(\gamma_3)$, and $H^2(x(\gamma_3))$ are given in Table II. They are calculated from series expansions for small γ_3 and from asymptotic formulas for large γ_3. Intermediate values, for which γ_3 is greater than 0.10, are somewhat uncertain; those for $F(\gamma_3)$ were found by numerical integration in a few cases; those for $\gamma_3F'(\gamma_3)$ were found from the first two differences of $F(\gamma_3)$.

When γ_3 is less than 0.04, we have to within five percent,

$$H^2(x)=4\ln\gamma_2/\gamma_1-0.688-6.7\gamma_1/\gamma_2. \quad (32)$$

Combining (17), (26b), and (32) we have, if γ_1 is less than $\gamma_2/10$ and γ_2 is less than 0.2,

$$H(x)=4\ln 1/\gamma_2-2.33+61\gamma_1+6.7\gamma_1/\gamma_2, \quad (33)$$

which is accurate to within 10 percent; thus $H(x)$ approaches asymptotically a value independent of γ_1. This asymptotic result holds also for β less than $\frac{1}{16}$, as we have seen above. When γ_1 is large, on the other hand, γ_3 will be even larger, since γ_2 is assumed to be small. For such values, $H^2(x)$ vanishes exponentially and $H(x)$ equals $H^1(x)$, which in this case is given by (26a).

These formulas give the profiles of a single component of a hydrogen line. For comparison with the corresponding result in the statistical case, we change from $J(x)$ to $W(\beta)$, defined by

$$W(\beta)d\beta=J(x)dx. \quad (34)$$

As we have seen—formula (13b)—β is the ratio of x to the static line shift produced by a stationary ion at a distance s; hence we have

$$\beta=xs^2h/q=\gamma_1/\gamma_2^2, \quad (35)$$

TABLE I. $H^1(x)$ *as a function of* γ_1*—formula (24).*

γ_1	$H^1(x)$	γ_1	$H^1(x)$	γ_1	$H^1(x)$
1.00	2.78	0.060	10.0	0.004	19.3
0.75	3.22	.050	10.6	.002	22.1
.50	3.94	.040	11.3	.001	24.7
.25	5.57	.030	12.3	.00075	25.8
.10	8.39	.020	13.6	.00050	27.4
.090	8.7	.010	15.9	.00025	30.2
.080	9.1	.008	16.7	.00010	33.8
.070	9.5	.006	17.8	.00001	43.0

where we assume for convenience that q is not a function of m; i.e., that all ions have the same charge. From (5), (8), (11), (34) and (35) we find

$$\beta^2 W(\beta) = 2J(x) \times x^2 s^2 \hbar/q, \tag{36a}$$

$$= 3\pi^{-\frac{1}{2}} \sum_m \gamma_2 H_m(x) N_m/N, \tag{36b}$$

where the first term in (8) has been assumed negligible. The right-hand side of (36) has been multiplied by two to make the result comparable with that of the usual statistical treatment, in which $J(x)$ is summed over only half the possible components of a spectral line, since each component will on the statistical theory be displaced wholly to the violet or wholly to the red, in general, depending on the sign of q.

In Fig. 2 is shown the function $\beta^2 W(\beta)$, computed from (36b) for various values of γ_2 on the assumption that only one type of perturbing ion is present. The solid line given for comparison is the statistical curve calculated from Holtsmark's[2] analysis by Verweij.[5] The dashed horizontal lines are the values derived from the Weisskopf[6] formula (22-I),

$$J(x) = \Omega_{1/\pi}/\pi x^2, \tag{37}$$

where $\Omega_{1/\pi}$ is the total number of collisions per second for which δ is less than $\delta_0(R')$, or in this case $1/\pi$. If we integrate $\Omega'(R, v)$ in (6) from zero to R', and determine R' from (2), then we find for (37), after we have integrated over v,

$$J(x) = \frac{2}{\pi^{\frac{1}{2}}} \sum_m \frac{\pi^2 q^2 l_m^{\frac{1}{2}} N_m}{\hbar^2 x^2}, \tag{38}$$

and therefore, from (36a),

$$\beta^2 W(\beta) = 3\pi^{\frac{1}{2}} \sum_m \gamma_2 N_m/N. \tag{39}$$

This result is simply (36b) with $H_m(x)$ set equal to π^2, a result which may be derived directly from (9), replacing $f_\sigma^2(\xi)$ by $2/\delta^2$ and integrating from $\delta = 1/\pi$ to $\delta = \infty$. Hence the true asymptotic value of $\beta^2 W(\beta)$ for small β differs from the Weisskopf value as π^2 differs from $4 \cdot \ln 1/\gamma_2 - 2.33$, as is evident from (33). This difference vanishes for $\gamma_2 = 0.047$. The excess of the Weisskopf formula for larger γ_2 is attributable to its neglect of the necessary substitution of $f_\sigma(\xi)$ for $f_\delta(\xi)$ in (1), while the opposite error for smaller γ_2 arises from the neglect, in the Weisskopf analysis, of the more distant collisions. These errors are in opposite directions; and for values of γ_2 between 0.3 and 0.017, the Weisskopf formula gives results for small β that are accurate to within 40 percent.

The profiles of Fig. 2 must now be summed over all components of a given hydrogen line. Let the subscript j denote a particular component. To evaluate the observable profile $J(x)$, or the atomic line absorption coefficient α_ν, one must form the weighted average of $J_j(x)$ over all the components. Combining (8) and (17), we find for $J(x)$, again normalized to unity,

$$J(x) = \frac{1}{2\pi x^2} \left\{ \sum_j \Gamma_j \varphi_j + 4 \frac{\pi^{\frac{1}{2}}}{x^{\frac{1}{2}} \hbar^{\frac{1}{2}}} \times \sum_j \varphi_j \sum_m N_m q_{mj}^{\frac{1}{2}} \gamma_1^{\frac{1}{2}} (H_{mj}^{1}(x) - H_{mj}^{2}(x)) \right\}, \tag{40}$$

where the φ_j are normalized oscillator strengths given by

$$\varphi_j = f_j / \sum_i f_j, \tag{41}$$

TABLE II. *$H^2(x)$ as a function of γ_3—formulas (28) to (31)*

γ_3	$F(\gamma_3)$	$-\gamma_3 F'(\gamma_3)$	$H^2(x)$
10^{-5}	9.88	0.995	25.1
10^{-4}	7.51	.982	20.5
0.001	5.28	.945	15.8
.010	3.18	.852	11.1
.020	2.62	.802	9.70
.040	2.08	.734	8.24
.060	1.80	.687	7.42
.080	1.61	.649	6.80
.10	1.45	.618	6.38
.20	1.20	.53	5.0
.40	0.71	.44	3.9
.60	.54	.38	3.2
.80	.44	.33	2.7
1.00	.37	.29	2.4
1.20	.31	.26	2.1
1.40	.27	.23	1.8
1.60	.24	.21	1.7
1.80	.22	.20	1.6
2.00	.198	.188	1.46
2.50	.160	.153	1.19
3.00	.136	.137	1.05
4.00	.102	.110	0.82
5.00	.076	.090	.66
6.00	.059	.076	.55
7.00	.048	.064	.46
8.00	.039	.055	.39
9.00	.033	.048	.33
10.0	.027	.042	.29
15.0	.014	.025	.172
20.0	.0084	.0158	.108
40.0	$2.1 \cdot 10^{-4}$	.0047	.030
80.0	$3.1 \cdot 10^{-5}$	.0010	.006

[5] S. Verweij, Publ. of the Astron. Inst. of the Univ. of Amst., No. 5 (1936).

[6] V. Weisskopf, Physik. Zeits. **34**, 1 (1933).

and where $H_{mj}^{1}(x)$ as a function of γ_1, and $H_{mj}^{2}(x)$ as a function of γ_3 (defined in (29)), are given in (24) and (28), and in Tables I and II, respectively. When γ_1 is less than one, it is convenient to make the substitution in (40)

$$\frac{q_{mj}^{\frac{1}{2}}\gamma_1^{\frac{1}{2}}}{x^{\frac{1}{2}}\hbar^{\frac{1}{2}}}=\frac{q_{mj}^{2}l_m^{\frac{1}{2}}}{\hbar^2}, \tag{42}$$

which follows from (11).

2

Apart from the physical assumptions made at the beginning of I, formula (40) should have an accuracy considerably greater than that provided by any previous impact theory. Three sets of approximation functions have been used: (14) and (15) for $f_\sigma(\xi)$ in the calculation of $L(x, u)$; (22a), (22b) and (23) for $L(x, u)$ in the calculation of $H^1(x)$; and finally (27) in the computation of $H^2(x)$. The first of these should produce an error in the smoothed $L(x, u)$ of less than 15 percent. The second and third are accurate, as we have already seen, to within 12 percent and 6 percent, respectively. Hence the total mathematical error cannot exceed 33 percent. If we take half the maximum error in each case as the probable error introduced from each source, the total probable error introduced by the mathematical approximations will be equal to or less than 9.7 percent. When γ_1 and γ_3 are both very different from unity, (40) should be accurate to within one or two percent.

The calculated profiles in Fig. 2 show two characteristics. Since a horizontal line represents a dispersion curve profile c/x^2, where c is some constant, the curves of Fig. 2 show in the first place that, as x increases, $J(x)$ changes from c/x^2 to the statistical profile, which in this case is $c'/x^{5/2}$. This transition from impact to statistical broadening was first pointed out on theoretical grounds by Margenau,[7] and has been observed in the laboratory as a transition in the rate of line growth which takes place with increasing pressure. Experimental difficulties have of course prevented a laboratory observation of this effect for the hydrogenic case analyzed here.

In the second place it is immediately evident

[7] H. Margenau, reference 3.

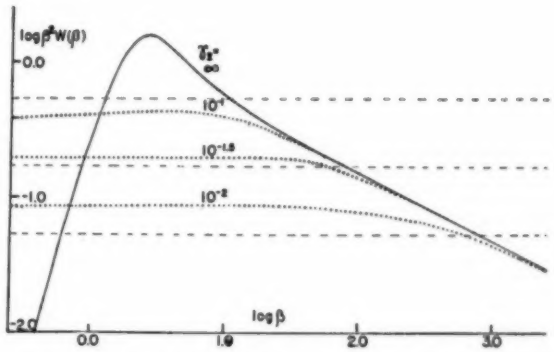


FIG. 2. Profile of a single component of a hydrogen line. Abscissae represent log β, where β is the ratio of $\nu-\nu_0$ to the displacement produced by a stationary ion at the average interionic distance s. Ordinates represent log $\beta^2W(\beta)$, where $W(\beta)d\beta$ is the probability of finding β in the range $d\beta$. The solid line is the Holtsmark statistical curve, while the dotted lines are calculated from (36b) for various values of γ_2, a quantity which varies inversely with the average ionic velocity. The horizontal dashed lines give the corresponding values of log $\beta^2W(\beta)$ derived from the Weisskopf impact formula.

from Fig. 2 that $J(x)$ decreases with γ_2 and hence decreases as the velocity increases. In an assembly containing equal numbers of electrons and positive ions, the latter are usually moving sufficiently slowly so that their value of γ_2 is greater than unity and their perturbations produce the statistical effect. The electrons may be moving sufficiently rapidly, however, so that their corresponding γ_2 is a very small quantity. Unless x is so large that the statistical formula holds for electrons as well as for positive ions, it follows that the contribution of the electrons to $J(x)$ will be negligible.

It may be noted that, subject again to the restrictions imposed in I, (40) includes all the relevant effects. "Collision broadening," regarded as producing transitions from one quantum level to another, is equivalent, not supplementary, to the above analysis.

The physical simplifications introduced in I may now be re-examined with the above analysis in mind. The assumption of single encounters is an adequate approximation when γ_2 is less than about a fifth. $J(x)$ in this case depends only logarithmically on γ_2, and because of this fact an arbitrary decrease in the effective s does not change the line profile appreciably, but increases γ_2 and eliminates the effect of distant multiple encounters. Multiple encounters at close distances have a negligible probability.

A second limitation of the theory is that R must be greater than $10\bar{r}_e$, or $7.91\cdot10^{-8}n_A^2$, where n_A is the total quantum number of the upper state of the transition in question, and $\bar{r}_e$ denotes the average value of r_e for the state $l=0$, $n=n_A$. This restriction is necessary primarily because of the inhomogeneity of the ionic field. One may find a first approximation to the error introduced by the neglect of this inhomogeneity. Let subscripts e denote electronic coordinates; R, the distance from the ion to the atomic nucleus; and ρ the distance from the electron to the ion; then

$$\rho^2=R^2+r_e^2-2z_eR, \tag{43}$$

where as usual the z_e axis is taken along R. If we let the nuclear charge be Ze, then the perturbing potential energy V is given by

$$V=-\frac{Ze^2}{\rho}=-\frac{Ze^2}{R}\left\{1+\frac{z_e}{R}+\frac{1}{2R^2}(3z_e^2-r_e^2)\cdots\right\}. \tag{44}$$

The term in z_e gives the usual Stark-effect displacement. To evaluate the next term we make use of the fact that it is the levels of large $\bar{z}_e$ that produce the observed line broadening; it is these levels, therefore, in which we are primarily interested. When z_e is large, the average value of r_e^2 is not much greater than the average value of z_e^2, which, in turn, is approximately given by $(\bar{z}_e)^2$. In such a case the perturbing potential (with neglect of the constant term) becomes roughly

$$\bar{V}=-\frac{Ze^2}{R^2}\bar{z}_e\left\{1+\frac{\bar{z}_e}{R}\right\}. \tag{45}$$

Since $\bar{z}_e$ is at most $\bar{r}_e$, the relative error in $\bar{V}$ produced in the most strongly shifted components by the inhomogeneity of the ionic field will be $\bar{r}_e/R$, to this approximation. This correction term will lead to an asymmetrical line, since it is independent of the sign of $\bar{z}_e$. This is reasonable physically, since a very close point charge will depress levels with positive $\bar{z}_e$ more than it will raise those of negative $\bar{z}_e$. When R is less than 3 or 4 times $\bar{r}_e$, this approximation may be expected to break down; higher terms in (44) and the higher order approximations of the perturbation theory will both become important.

To examine more closely the range of validity of the theory, and to clarify the meaning of the quantities used in Section 1, we derive several useful relationships. From (11-I) and (12-I) we see that the Stark-effect frequency shift $\Delta\nu_{AB}$ of a given component produced by a stationary charge Z_me at a distance R is given by

$$\Delta\nu_{AB}=\frac{q_m}{hR^2}=\frac{3\hbar Z_m}{4\pi m_e}\frac{g}{R^2}=\frac{0.275gZ}{R^2}, \tag{46}$$

$$g=n_A(k_{2A}-k_{1A})-n_B(k_{2B}-k_{1B}), \tag{47}$$

where k_1 and k_2 are the quantum numbers for the ξ and η coordinates, respectively. If Z_m is taken to be unity, R is taken as the distance of closest approach, and g is replaced by pn_A^2, where $|p|$ ranges from zero to approximately unity, one finds from (2), (46), (7) and (11) that

$$\delta=\frac{1.73pn_A^2}{vR}=\frac{2.19\cdot10^8p}{v}\frac{\bar{r}_e}{R}, \tag{48}$$

$$\Delta\lambda_{AB}=1.22\cdot10^3\left(\frac{\bar{r}_e}{R}\right)^2\frac{pn_B^4}{n_A^2[1-(n_B/n_A)^2]^2}A, \tag{49}$$

$$\gamma_1=2.18\cdot10^8p\frac{w_m}{T}\frac{\Delta\lambda}{\lambda_0}\left\{\left(\frac{n_A}{n_B}\right)^2-1\right\}, \tag{50a}$$

$$\gamma_2^2=4.74\cdot10^{-8}p^2\frac{w_m}{T}n_A^4N_i^2, \tag{50b}$$

where A denotes angstrom units, T the absolute temperature, and w_m is defined as $M_m/(M_m+m_H)$; for an electron, w_m is $5.4\cdot10^{-4}$, for a proton it is $\frac{1}{2}$, and for other ions it is practically unity; $2\pi\Delta\nu_{AB}$ is simply $x_{\max}$, $\Delta\lambda_{AB}$ equals $C\Delta\nu_{AB}/\nu_0^2$, and $\Delta\lambda$ equal $cx/2\pi\nu_0^2$. It is clear from (48) that when R is as small as $10\bar{r}_e$, δ will be very large unless v is at least 10^6 cm/sec. This is an observable thermal velocity only in the case of electrons, and for heavy atoms such close encounters will accordingly be adiabatic.

For these adiabatic encounters the statistical theory is valid and the wings of a line will be produced only by very close encounters. In this case $J(x)$ is proportional to $R^2(x)dR/dx$; $4\pi R^2(x)dR$ is the volume of the shell from which a stationary ion produces a Stark-effect shift between x and $x+dx$ in angular frequency units. To a first approximation dR/dx is unchanged by the inhomogeneity, provided that x is held fixed, and hence the relative error in $J(x)$ is simply the relative error in $R^2(x)$ for a known x, or approxi-

mately the relative error in x for a given R, since x is approximately inversely proportional to R^2. If we assume that the broadening of the spectral line is produced largely in the upper state of the transition, then it follows from (45) that the correction factor for $J(x)$ in this case is approximately $1\mp\bar{r}_e/R$, where the minus sign goes with positive x and vice versa; this assumes that the wings of the line are produced by the most strongly shifted Stark-effect components.

When R is less than $3\bar{r}_e$ the theory is no longer valid; hence from (49) it is evident that the formulas for $J(x)$ do not apply if $\Delta\lambda_{AB}$ for the Balmer series ($n_B=2$) exceeds $2160/n_A^2$ angstroms. If n_A is greater than 7, this upper limit is greater than half the separation between successive lines, and is hence not a serious restriction. For the lower members of the series this limiting shift is greater than is usually observed in stellar spectra.

When electrons are involved, the case is somewhat different. At a temperature of 5000° the electronic velocity is $4\cdot10^7$ cm/sec., and δ is only $\frac{1}{2}$ when R equals $10\bar{r}_e$. A substantial portion of $L(x, u)$ and of $H(x)$ arises from encounters for which δ is greater than this value; unlike the statistical case discussed above, the inhomogeneity of the ionic field at close distances will therefore affect $J(x)$ even when x is relatively small. A rigorous treatment would be difficult here, but it is not difficult to set an upper limit on the error involved provided that γ_1 is small.

In such a case $L(x, u)$ may be regarded as the sum of two parts—the first comprising all encounters for which δ is greater than some critical value δ_0, which we may take equal to 1, the second including the more distant encounters. The first part follows approximately the Weisskopf formula (37), where now unity replaces $1/\pi$ as the critical value of the phase shift. The second part, which is six times as great as the first provided γ_1 is less than $\frac{1}{3}$, and twenty times as great if γ_1 is less than 1/30, arises from the more distant encounters, and will be largely unaffected by the inhomogeneity of the perturbing field. The relative error in δ will be $2\bar{r}_e/\pi R$, as we see from (45); the extra factor $2/\pi$ arises from the integration of $\bar{V}(t)$ over t, replacing R by $(v^2t^2+R^2)^{\frac{1}{2}}$ as in I. When R equals $3\bar{r}_e$, this gives a relative error of 21 percent in R, for a fixed δ, and, since the total number of relevant collisions is proportional to R^2, a relative error of 42 percent in the first part of $L(x, u)$. This gives a corresponding error of less than six percent in the sum of both parts, if γ_1 is less than $\frac{1}{3}$. If δ is less than unity when R equals $3\bar{r}_e$, and hence if the temperature exceeds 14,000°, this error must be increased slightly. The relative increase in the error will presumably be less than the relative increase in T, however, and hence if T is as much as 30,000°, the relative error should not exceed 13 percent. As γ_1 increases the error rises until it approaches the corresponding error in the statistical case, which according to (45) may be as much as 30 percent.

One may conclude, then, that (40), supplemented by the Holtsmark-Verweij statistical formula, gives values of $J(x)$ which, if one neglects the inhomogeneity of the ionic field, have a probable error of not more than ten percent and are asymptotically correct, except for the intractable case when γ_2 is approximately between 0.2 and 2 and γ_1 is at the same time less than unity. The inhomogeneity introduces a correction factor for the statistical case; for the impact case—when γ_1 is small—the additional error introduced is less than 13 percent, provided T does not exceed 30,000°. If, as before, we take half the upper limit on the error as the value of the probable error, this gives a total probable error for the impact case of less than 12 percent. When $\Delta\lambda_{AB}$ is greater than $135n_B^4/n_A^2$ angstroms, the theory breaks down.

One more restriction may be noted. The neglect of the cross-product terms in (16-I) and (1) is legitimate only if x is sufficiently far removed from the line center. It is of course physically obvious that if the line is very much widened by perturbations, the normalized intensity must decrease over a central region whose width may be many times greater than the value of the radiative damping constant Γ.

It is a very real pleasure to record here my indebtedness to the National Research Council for a grant which has made this work possible, and to Dr. Harlow Shapley, Dr. Donald Menzel, and the staff of the Harvard College Observatory for the opportunity to use the facilities of the Observatory and to profit by many informal and informative discussions.

The Zeeman Effect of the Hyperfine Structure of Optically Excited Mercury Resonance Radiation

E. HOBART COLLINS*
Parsons College, Fairfield, Iowa
(Received April 22, 1939)

Magnetic fields from 0 to 2000 gauss are applied to the mercury resonance line λ2537A and the Zeeman patterns for the hyperfine structure determined. The results indicate close agreement with the explanation of Schuler and Keyston. The Paschen-Back effects for the hyperfine structure of the isotopes 199 and 201 are worked out and experimental evidence is offered in support of the hypothesis of Goudsmit and Bacher that the Zeeman effect for the hyperfine structure may be treated in the same manner as the Zeeman effect for the fine structure if the appropriate change in quantum numbers is made.

INTRODUCTION

IF THE explanation of the hyperfine structure of mercury resonance radiation, λ2537A, $1^1S_0-2^3P_1$, offered by Schuler and Keyston[1] is accepted, certain predictions can be made as to the expected polarization and also as to the behavior of the components in a magnetic field. Larrick and Heydenburg[2] have studied the polarization and have also restated the explanation of Schuler and Keyston. It is the purpose of this paper to report on an investigation of the effect of applying magnetic fields to optically excited mercury resonance radiation. From the viewpoint of hyperfine structure these fields would be considered strong inasmuch as complete overlapping of the components was obtained. However, from the standpoint of fine structure they are very weak fields. Optically excited radiation was chosen because of the necessity of using a source of light having the smallest possible breadth of spectral line.

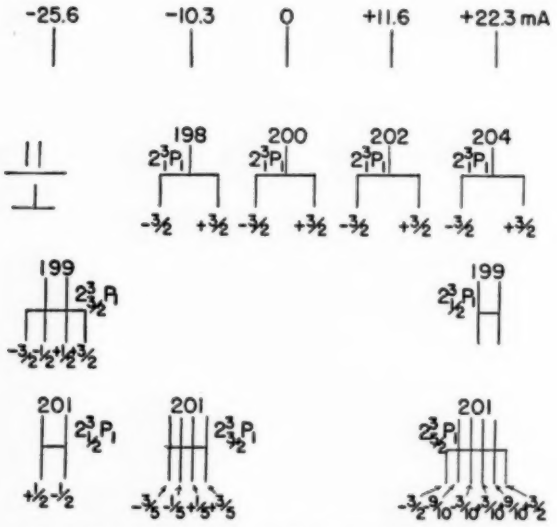


FIG. 1. Parallel and perpendicular splittings computed on the basis of Schuler and Keyston's assignment of quantum numbers and nuclear magnetic moments.

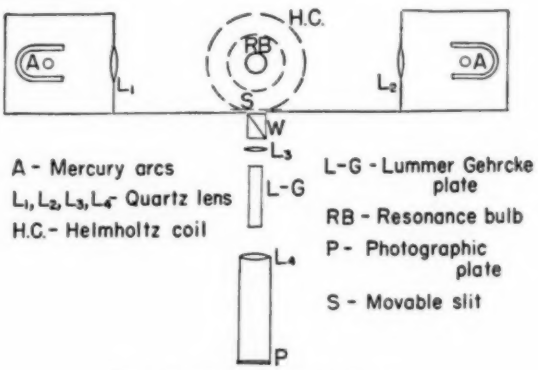


FIG. 2. Diagram of apparatus.

Figure 1[3] shows the parallel and perpendicular splittings computed on the basis of Schuler and Keyston's assignment of quantum numbers and nuclear magnetic moments. To calculate these splittings one computes the Landé hyperfine structure factor,[4] (g_f), for the upper 2^3P_1 level.

* The investigations were performed during summer appointments at The State University of Iowa, Iowa City, Iowa.

[1] H. Schuler and J. E. Keyston, Naturwiss. **31**, 676 (1931).

[2] L. Larrick and N. P. Heydenburg, Phys. Rev. **39**, 289 (1932).

[3] The notation used in all the figures and in the discussion is that used by H. E. White, *Introduction to Atomic Spectra* (McGraw-Hill, 1934), Chapter 18.

[4] Equation 18.52, page 374, H. E. White, *Introduction to Atomic Spectra*.

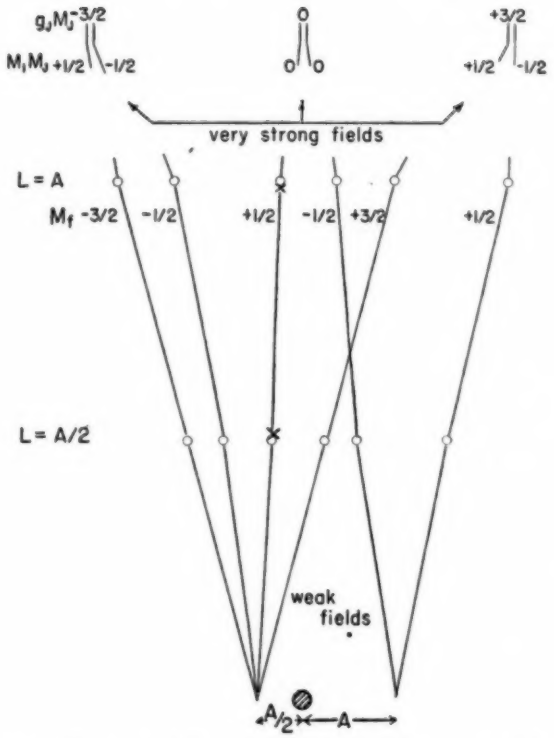


FIG. 3. Theoretical Zeeman splittings to be expected for the odd-numbered Hg isotope, 199.

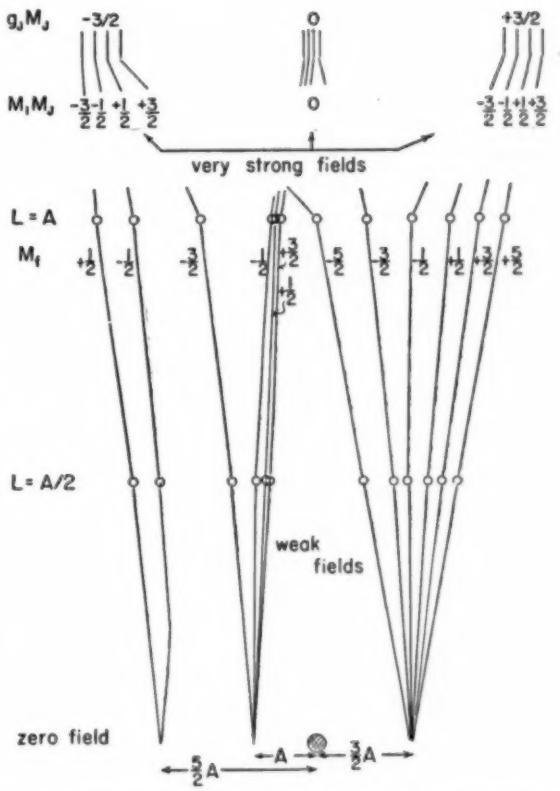

$g_J M_J$ -3/2 0 +3/2
$M_1 M_J$ $-\frac{3}{2}$ $-\frac{1}{2}$ $+\frac{1}{2}$ $+\frac{3}{2}$ 0 $-\frac{3}{2}$ $-\frac{1}{2}$ $+\frac{1}{2}$ $+\frac{3}{2}$
very strong fields
L = A
M_f $+\frac{1}{2}$ $-\frac{1}{2}$ $-\frac{3}{2}$ $-\frac{1}{2}$ $+\frac{3}{2}$ $+\frac{1}{2}$ $-\frac{5}{2}$ $-\frac{3}{2}$ $-\frac{1}{2}$ $+\frac{1}{2}$ $+\frac{3}{2}$ $+\frac{5}{2}$
L = A/2
weak fields
zero field
$\frac{5}{2}A$ A $\frac{3}{2}A$

FIG. 4. Theoretical Zeeman splittings to be expected for the odd-numbered Hg isotope, 201.

Since the lower level, 1^1S_0, does not split the theoretical splitting of the lines themselves is that of the upper level.

APPARATUS

A diagram of the apparatus is shown in Fig. 2. The source of radiation for the optical train through the Lummer-Gehrecke interferometer to the photographic plate was within the resonance bulb. This resonance bulb had a tube extending downward about 25 cm containing a drop of mercury. The vapor pressure in the bulb was reduced by maintaining the tube at 0°C, during the time that pictures were being taken. The magnetic field was produced by a Helmholtz coil made of copper tubing carrying a stream of water to remove excess heat. The coil produced a magnetic field of 5.46 gauss per ampere and was capable of carrying a current of 600 amperes. The mercury arcs used to excite the radiation in the resonance bulb were water cooled to decrease the width of the spectral line and magnetically deflected to eliminate self reversal. The Lummer-Gehrecke plate was 4.92 mm thick so that the order separation is 45.1mA and produced a pattern previously described by Ellett and MacNair.[5] Unfortunately this order separation is such that the +22.1 and the −25.6 milliangstrom components overlap making the magnetic analysis of these components difficult. The slit, Wollaston prism, and the quartz lens, L_3, produced two slightly diverging, approximately parallel beams of light, one polarized parallel to, the other perpendicular to the magnetic field. The camera lens focused the interferometer pattern of the Lummer-Gehrecke plate superimposed on the double image of the slit. A slight sidewise change in the position of the slit made it possible to secure on the same photographic plate patterns of two different fields. This proved to be very useful in establishing the identity of the various components as the field changed from one value to the next. The Lummer-Gehrecke plate and camera were care-

[5] A. Ellett and W. A. MacNair, Phys. Rev. **31**, 180 (1928).

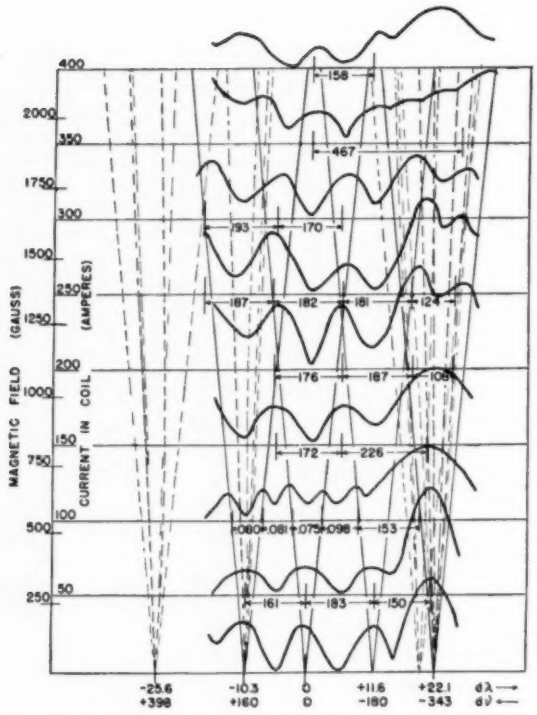


FIG. 5. Zeeman separations for the perpendicular components of Hg resonance radiation. Full lines for even-numbered isotopes. Dotted lines for odd-numbered.

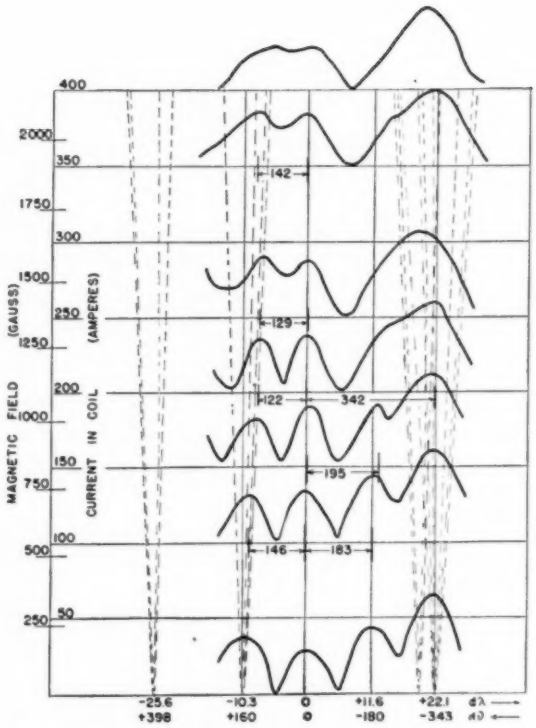


FIG. 6. Zeeman separations for the parallel components of Hg resonance radiation. Full lines for even-numbered isotopes. Dotted lines for odd-numbered.

fully shielded from scattered light and excess heat. The microphotometer traces of each of the plates used in each of the measurements were obtained with a high precision microphotometer.

THE PASCHEN-BACK EFFECT

Goudsmit and Bacher[6] have pointed out that the formulae developed by Darwin for the Zeeman effect as applied to fine structure will also apply to the hyperfine structure if the appropriate change in quantum numbers is made. On the basis of these ideas the theoretical splittings to be expected for the odd isotopes has been worked out and is shown graphically in Figs. 3 and 4. Attention is called to the fact that in Fig. 3 the component with the displacement $\frac{1}{2}A$ and $M_I=+\frac{1}{2}$ is evidently the component which MacNair has reported to be displaced from the zero component in accord with the equation $d\lambda=-25.6+0.089\,H^{\frac{1}{2}}$, Darwin's theory gives $E=-\frac{1}{4}A-\frac{3}{4}L+\frac{1}{4}(9A^2+6AL+9L^2)^{\frac{1}{2}}$ with the displacement measured from the center of gravity of the line. At $L=A$ and $L=\frac{1}{2}A$ the two equations, if the appropriate change in units is made, are in agreement to within 1 percent. In Fig. 3, two crosses indicate this agreement graphically. Because of the overlapping of the two components $+22.1$ and -25.6 milliangstrom units in the Lummer-Gehrecke patterns the author was unable to observe this shift experimentally. In Fig. 4, the zero field component with a displacement A to the left of center of gravity breaks into 4 Zeeman splittings, three of which shift toward the center of gravity. Experimentally it was observed (see Fig. 6) that for the parallel components there is very good agreement with this theoretical displacement.

RESULTS

The microphotometer traces were measured to obtain the $d\lambda$ separation in milliangstrom units and the corresponding $d\nu$ in reciprocal centimeters. The height of the peaks above the general blackening of the photographic plate was taken as a rough measure of intensity of the

[6] S. Goudsmit and R. F. Bacher, Phys. Rev. **34**, 1499 (1929).

components. These measurements enable the sketching in of the microphotometer traces with intensity as ordinates and position of the components as abscissae. These traces are placed one above the other to show the progressive change in the pattern as magnetic fields of greater and greater strength were applied. The results for the perpendicular components are shown in Fig. 5 and the parallel components in Fig. 6. In both of these figures the vertical lines or the lines at a small angle from the vertical represent the theoretical splitting for the even-numbered isotopes as shown in Fig. 1 and for the odd-numbered isotopes as shown in Figs. 3 and 4. The splittings for the even-numbered and the odd-numbered isotopes are shown as full and dotted lines, respectively. The perpendicular components show $\frac{3}{2}$ splitting which is to be expected on the basis of Schuler and Keyston's explanation and which has previously been reported by MacNair.[5] The examination of the experimental results for the parallel components show that the -10.3 component shifts toward the zero component in agreement with the theoretical prediction of Darwin's theory but which for the range of fields used is not in agreement with the results of MacNair.

ACKNOWLEDGMENTS

The writer wishes to thank the National Research Council for a grant-in-aid for a preceding research of which this investigation is the direct result, also The State University of Iowa for an honorarium during the summer of 1938 when the work was completed. He wishes to express his appreciation to Professor A. Ellett who suggested the problem and who gave excellent guidance during the period of investigation.

JULY 1, 1939 PHYSICAL REVIEW VOLUME 56

The Paschen-Back Effect

V. Theory of the Effect for Intermediate Coupling

J. B. GREEN AND J. F. EICHELBERGER
Mendenhall Laboratory of Physics, Ohio State University, Columbus, Ohio

(Received April 24, 1939)

General spectroscopic theory has been applied to the Paschen-Back effect and simplified methods developed for determining the positions of the energy levels and intensities of lines for cases of intermediate coupling. They are expressed in terms of *LS*-coupling as the zero-order functions.

A COMPLETE theory of the Paschen-Back effect for intermediate coupling in two-electron spectra was first given by Houston[1] (in essence) for the case of one electron in an *s* state and was later completed and experimentally verified by Green and Loring[2] and by Jacquinot.[3] Houston's zero-order approximation, while satisfactory for the particular example studied, proves unwieldy, however, when applied to more complicated configurations. Since the publication of Houston's paper, a very considerable body of theoretical work has been carried out, and spectroscopic theory has been advanced to the point where the problem of the Paschen-Back effect can be stated in general terms for any kind of coupling. While it is useful to be able to state the general solution to any problem, it is the actual application of the problem to particular cases that is of prime interest to the experimental physicist; and it is the purpose of the present paper to express the solution in terms that are very easily interpreted in terms of a coupling system that is familiar to everybody; namely, *LS*-coupling.

[1] W. V. Houston, Phys. Rev. **33**, 297 (1929).
[2] J. B. Green and R. A. Loring, Phys. Rev. **46**, 888 (1934).
[3] P. Jacquinot, Thesis, (Paris, 1937).

I. Calculation of Energy Levels

Let us suppose that using the LS-coupling scheme we can write part of the matrix of a given configuration involving the two j values j and $j+1$, in the following manner.

	s_{L_j}	$s^{-1}L_j$	$s_{(L-1)_j}$	$s_{L_{j+1}}$	$s_{(L+1)_{j+1}}$	$s_{(L-1)_{j+1}}$
s_{L_j}	A	a	b			
$s^{-1}L_j$	a	B	c		0	
$s^{-1}(L-1)_j$	b	c	C			
$s_{L_{j+1}}$				D	d	e
$s_{(L+1)_{j+1}}$		0		d	E	f
$s_{(L-1)_{j+1}}$				e	f	F

(1a)

where the lower case letters represent magnetic interaction terms arising from spin-orbit interaction, and the upper case letters represent, in general, electrostatic plus magnetic (spin-orbit) interaction energies. In a large number of cases, especially for two-electron spectra, these terms have been calculated[4] in terms of certain parameters representing radial integrals. It is possible, then, from the known energy levels of a given spectrum, to work backwards and determine the values of the parameters. Let us suppose that the observed energy levels are A', B', C', D', E', F'. This is equivalent to saying that we have determined a transformation.

	s_{L_j}	$s^{-1}L_j$	$s_{(L-1)_j}$	$s_{L_{j+1}}$	$s_{(L+1)_{j+1}}$	$s_{(L-1)_{j+1}}$
A_j'	a_1	b_1	c_1			
B_j'	a_2	b_2	c_2		0	
C_j'	a_3	b_3	c_3			
D_{j+1}'				d_1	e_1	f_1
E_{j+1}'		0		d_2	e_2	f_2
F_{j+1}'				d_3	e_3	f_3

(2a)

which transforms the matrix (1) into a diagonal matrix. These transformation coefficients are normalized and orthogonal, i.e., if $a_1^2+a_2^2+a_3^2=1$, etc. and $a_1a_2+b_1b_2+c_1c_2=0$, etc.

In the presence of a magnetic field, there is, in addition to the matrix of the energies represented by (1), a matrix resulting from the interaction of L and S with the magnetic field, which may be represented for each value of the magnetic quantum number m, in the following manner

	s_{L_j}	$s^{-1}L_j$	$s_{(L-1)_j}$	$s_{L_{j+1}}$	$s_{(L+1)_{j+1}}$	$s_{(L-1)_{j+1}}$
s_{L_j}	$mg_1\omega$	0	0	$k\omega$	0	0
$s^{-1}L_j$	0	$mg_2\omega$	0	0	0	0
$s_{(L-1)_j}$	0	0	$mg_3\omega$	0	0	$l\omega$
$s_{L_{j+1}}$	$k\omega$	0	0	$mg_4\omega$	0	0
$s_{(L+1)_{j+1}}$	0	0	0	0	$mg_5\omega$	0
$s_{(L-1)_{j+1}}$	0	0	$l\omega$	0	0	$mg_6\omega$

(3a)

where $\omega=ehH/4\pi\mu_0c$, and the g_i's are the Landé g-factors

$$g_i=1+\frac{J(J+1)-L(L+1)+S(S+1)}{2J(J+1)},$$

$$k \text{ and } l=\left[\frac{(J-L+S+1)(J+L-S+1)(J+L+S+2)(L+S-J)}{4(J+1)^2(2J+1)(2J+3)}\right]^{\frac{1}{2}}[(J+1)^2-m^2]^{\frac{1}{2}}$$

and are diagonal, while nondiagonal terms occur only between terms with J's differing by unity but with the same S and L.

If we apply the transformation (2) to this matrix, and add the result to the diagonalized matrix in the absence of the magnetic field, the problem of the determination of the energy levels is complete.

To make the case simple, suppose that three of the levels, A', B', D' are sufficiently close so that they will perturb each other in the presence of a magnetic field. The part of the final matrix which then becomes important for the calculations of A'', B'', D'' (the altered positions of A', B', D' in a strong magnetic field) is

	$A'_{j,m}$	$B'_{j,m}$	$D'_{j+1,m}$
$A'_{j,m}$	$A'+mg_1'\omega$	$mv\omega$	$w\omega$
$B'_{j,m}$	$mv\omega$	$B'mg_2'\omega$	$y\omega$
$D'_{j+1,m}$	$w\omega$	$y\omega$	$D'+mg_3'\omega$

(4a)

(C', E', F' are so distant that their effects on A'', B'', D'' are negligible), where $g_i'=a_i^2g_1$

[4] See e.g., E. U. Condon, and G. H. Shortley, *Atomic Spectra* (Cambridge University Press, 1934).

$+b_i^2g_2+c_i^2g_3$, which represents the weak-field magnetic splitting factor in the intermediate coupling,

$$v=a_1a_2g_1+b_1b_2g_2+c_1c_2g_3$$
$$w=a_1d_1k+c_1f_1l$$
$$y=a_2d_1k+c_2f_1l.$$

If we now subrract E from each of the diagonal terms of (4) and set the determinant of this matrix equal to zero, the three roots will represent the positions of the energy levels when a magnetic field is present. One noteworthy fact now becomes evident from (4). In LS-coupling there is no nondiagonal element between two terms with the same value of J. In intermediate coupling, these terms may be as large as any of the other terms introduced by the magnetic field.

II. The Intensities

The calculation of the intensities of the components involved in a transition between the configuration discussed above and another configuration whose transformation matrix is given below (5), is a much more tedious process

	$^sL_{j+1}$	$^s(L+1)_{j+1}$	$^{s-1}(L-1)_{j+1}$	$^s(L-1)_{j+1}$
H_{j+1}'	h_1	k_1	l_1	n_1
K_{j+1}'	h_2	k_2	l_2	n_2
L_{j+1}'	h_3	k_3	l_3	n_3
N_{j+1}'	h_4	k_4	l_4	n_4

(5a)

(In the following example the above configuration is spread out sufficiently so that no Paschen-Back effect in this configuration need be considered.)

If we neglect the quantities depending on the numbers of atoms in the different states and the factors depending on the frequency, then the matrix of the square roots of the line strengths in a particular transition may be written as in (6). (This procedure is justifiable, since we are usually dealing with an extremely small frequency interval in the Paschen-Back effect.) (6) represents the (strengths)$^{\frac{1}{2}}$ of transitions between levels j, m and $j+1$, m of one configuration, and $j+1$, $m+1$ of another configuration.

	$^sL_{j+1,\,m+1}$	$^s(L+1)_{j+1,\,m+1}$	$^{s-1}(L-1)_{j+1,\,m+1}$	$^s(L-1)_{j+1,\,m+1}$
$^sL_{j,\,m}$	p	q	0	t
$^{s-1}L_{j,\,m}$	0	0	r	0
$^s(L-1)_{j,\,m}$	s	0	0	u
$^sL_{j+1,\,m}$	α	β	0	γ
$^s(L+1)_{j+1,\,m}$	δ	ϵ	0	0
$^s(L-1)_{j+1,\,m}$	ζ	0	0	ξ

(6a)

Operating on (6) on the left by (2) and on the right by (5) transposed we obtain the matrix of $S^{\frac{1}{2}}$ in the type of coupling involved, in the absence of the magnetic field. The individual matrix elements are of the following form:

$$\begin{aligned}(A'jm\,|\,S^{\frac{1}{2}}\,|\,H'\,j+1\;m+1)&=(a_1p+c_1s)h_1\\&\quad+a_1qk_1+b_1rl_1+(a_1t+c_1u)n_1\\(E'j+1\;m\,|\,S^{\frac{1}{2}}\,|\,L'\,j+1\;m+1)&\\=(d_2\alpha+e_2\delta+f_2\zeta)h_3+(d_2\beta+e_2\epsilon)k_3&\\+0\cdot l_3+(d_2\gamma+f_2\xi)n_3.&\end{aligned}\qquad(7a)$$

The squares of these quantities would represent the line-strengths in a weak magnetic field.

If, now, a strong magnetic field is introduced it is necessary to diagonalize (4) before we can calculate the matrix of $S^{\frac{1}{2}}$. With the same approximations as in (4) the transformation matrix is

	$A'_{j,\,m}$	$B'_{j,\,m}$	$D'_{j+1,\,m}$
A''_m	a_1'	a_2'	a_3'
B''_m	b_1'	b_2'	b_3'
D''_m	d_1'	d_2'	d_3'

(8a)

(To be exact, we should include terms a_4', a_5', a_6', etc. due to interactions between A', B', D', and C', E', F', but these are negligible in the case considered.) Then the line-strength of the component for $A''_{j,\,m}\rightarrow H''_{j+1,\,m+1}$ (the values of A' and H' in a strong magnetic field) is given (approximately) by

$$\begin{aligned}&[(A''j\,m\,|\,S^{\frac{1}{2}}\,|\,H''\,j+1\;m+1)]^2\\&=[a_1'\{(a_1p+c_1s)h_1+a_1qk_1+b_1rl_1\\&\quad+(a_1t+c_1u)n_1\}+a_2'\{(a_2p+c_2s)h_1\\&\quad+a_2qk_1+b_2rl_1+(a_2t+c_2u)n_1\}\\&\quad+a_3'\{(d_1\alpha+e_1\delta+f_1\zeta)h_1\\&\quad+(d_1\beta+e_1\epsilon)k_1+0\cdot l_1+(d_1\gamma+f_1\xi)n_1\}]^2\end{aligned}\qquad(9a)$$

and similarly for the other components.

The authors are indebted to Dr. G. H. Shortley for many helpful discussions.

The Paschen-Back Effect

VI. The Spectrum of Neon

J. B. GREEN, *Mendenhall Laboratory of Physics, Ohio State University, Columbus, Ohio*

AND

J. A. PEOPLES, JR., *Lehigh University, Bethlehem, Pennsylvania*

(Received April 24, 1939)

The methods developed in the previous paper have been applied to the spectrum of neon. The transformation coefficients in this spectrum can be easily determined since parameters have been calculated for several of its configurations. When applied to several lines of the p^5p-p^5s and p^5p-p^5d transitions, results are obtained which are in extremely good agreement with observed patterns.

SOME observations on the Paschen-Back effect in neon have been given by Jacquinot[1] for the transitions $2p^53p\rightarrow 2p^5ms$ making use of Houston's[2] calculations for the ps configuration. The present paper is a discussion of transitions between $2p^53s\rightarrow 2p^54p$ and $2p^53p\rightarrow 2p^5md$, with the methods developed in the previous paper.[3]

The apparatus used has already been described.[4] In fact, the plates had been used for measurements on the Zeeman effect in neon.

THE TRANSITIONS $2p^54p\rightarrow 2p^53s$

The configuration $2p^53s$ has been shown by Shortley[5] to have the following eigenfunctions:

$$\begin{aligned}\psi(s_2)&=0.266\psi(^3P_1)-0.964\psi(^1P_1)\\ \psi(s_4)&=0.964\psi(^3P_1)+0.266\psi(^1P_1),\end{aligned}\tag{1b}$$

while s_3 is $^3P_0$ and s_5 is $^3P_2$. Bartberger[6] has determined the parameters of the $2p^54p$ configuration and from these may be calculated the transformation matrix. Of the ten levels of this configuration, the levels p_2 and p_4 are only 0.90 cm^{-1} apart and therefore produce large perturbations in the magnetic field. Their eigenfunctions in the absence of a magnetic field are

$$\begin{aligned}\psi(p_2)&=-0.055\psi(^3D_1)+0.789\psi(^3P_1)\\ &\quad+0.511\psi(^1P_1)-0.339\psi(^3S_1)\\ \psi(p_4)&=0.607\psi(^3D_2)+0.595\psi(^1D_2)\\ &\quad-0.528\psi(^3P_2).\end{aligned}\tag{2b}$$

These two levels perturb each other, therefore, because they both "consist" partially of ^3D and ^3P. The nondiagonal matrix element between them for $m=1$ would then be given by (see matrix (4a), previous paper)

$$\{(9/20)^{\frac{1}{2}}\times-0.055\times0.607+\tfrac{1}{2}\times0.789\times-0.528\}\omega=-0.230\omega,$$

while $g(p_4)=1.201$ and $g(p_2)=1.425$.

For a typical field strength, 29,000 gauss, corresponding to $\omega=1.367$ cm^{-1} the matrix for $m=1$ becomes

	p_4	p_2
p_4	1.64	−0.31
p_2	−0.31	−0.90+1.95

whose determinant has roots 0.91 and 1.77 indicating a shift of 0.13 cm^{-1} from the unperturbed positions. The perturbed eigenfunctions then become, by inserting the eigenvalues in the determinant of the matrix,

$$\begin{aligned}\psi(p_4')&=0.919\psi(p_4)-0.395\psi(p_2)\\ \psi(p_2')&=0.395\psi(p_4)+0.919\psi(p_2).\end{aligned}\tag{3b}$$

These values, together with (2) and (1) enable us to find the intensities of the components.

The calculation of the strength of a typical component would then proceed as follows. First, it is necessary to determine the intensities of the transitions in LS-coupling. Goldberg[7] has given these for a large number of transitions. In particular, the relative strengths of the multi-

[1] P. Jacquinot, Diss. Paris, 1937.
[2] W. V. Houston, Phys. Rev. **33**, 297 (1929).
[3] J. B. Green and J. F. Eichelberger, Phys. Rev. **56**, 51 (1939).
[4] J. B. Green and J. A. Peoples, Phys. Rev. **54**, 602 (1938).
[5] G. H. Shortley, Phys. Rev. **44**, 666 (1933).
[6] C. L. Bartberger, Phys. Rev. **48**, 682 (1935).
[7] L. Goldberg, Astrophys. J. **82**, 1 (1935).

plets in the transition $p^5s \rightarrow p^5p$ is given by

	^3D	^3P	^3S			
^3P	150	90	30	^1D	^1P	^1S
			^1P	50	30	10

Each of these numbers represents the sum of all of the strengths of the particular multiplet. Within the multiplet say $^3D^3P$ the strengths of the individual components are

		^3D		
		1	2	3
	0	16.6		
^3P	1	12.5	37.5	
	2	0.8	12.5	70.0

while the strengths of the components of the transition in a weak magnetic field of $^3D_2$ $^3P_1$ are given by

		$^3D_2$				
	$m=-2$	-1	0	1	2	
	1			0.9	5.6	*5.6*
$^3P_1$	0		2.8	8.4	*2.8*	
	-1	5.6	5.6	*0.9*		

(4b)

In the previous paper, (9a) gives the expression for a particular Paschen-Back component in terms of the square-roots of the quantities in (4b). But it is very important that the phases of these square roots be taken correctly. In (4b), the numbers in italics indicate that the negative square root is to be chosen. All the others use the positive square roots. Table I, condensed from Condon and Shortley,[8] gives the rules for choice of sign.

Then by 9(b), the intensity

$$\{(p_4'\ 2\ 1|S^{\frac{1}{2}}|s_4\ 1\ 0)\}^2=\{0.919[0.607\times0.964 \times-(2.8)^{\frac{1}{2}}+0.595\times0.266\times-(3.75)^{\frac{1}{2}}-0.528 \times0.964\times-(0.95)^{\frac{1}{2}}]-.395[-0.055\times0.964 \times(1.5)^{\frac{1}{2}}+0.789\times0.964\times(0.95)^{\frac{1}{2}}+0.511 \times0.266\times(3.75)^{\frac{1}{2}}-0.399\times0.964\times(1.25)^{\frac{1}{2}}]\}^2$$

[8] E. U. Condon and G. H. Shortley, *Theory of Atomic Spectra* (Cambridge University Press, 1935), p. 193.

TABLE I.

		L $L+1$	L L	L $L-1$
$J\rightarrow J+1$	$m\rightarrow m\pm1$	$\mp$	$\pm$	$\pm$
	$m\rightarrow m$	$+$	$-$	$-$
$J\rightarrow J$	$m\rightarrow m\pm1$	$+$	$+$	$+$
	$m\rightarrow m$	(sign of m)	(sign of m)	(sign of m)
$J\rightarrow J-1$	$m\rightarrow m\pm1$	$\mp$	$\mp$	$\pm$
	$m\rightarrow m$	$-$	$-$	$+$

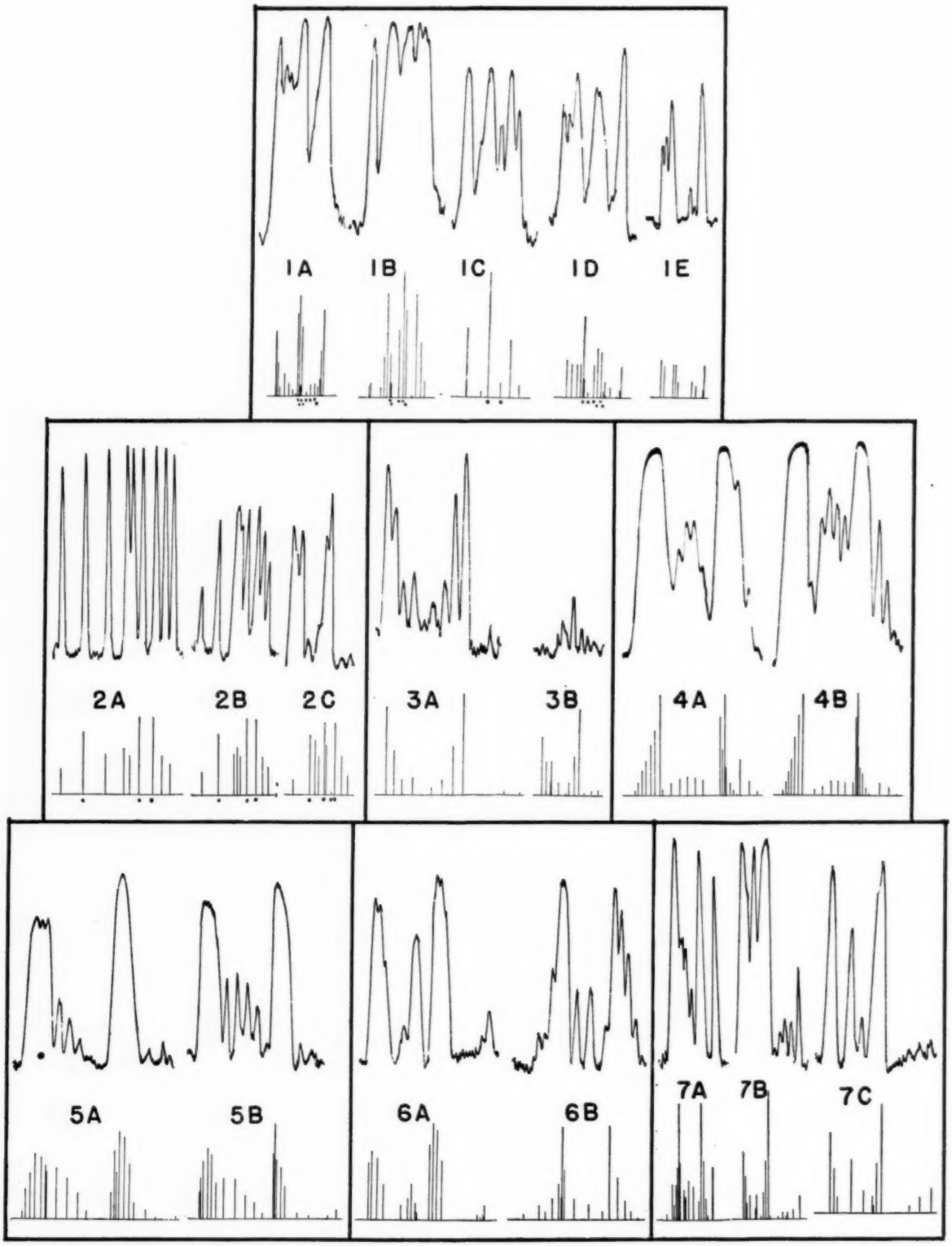
and similarly for other components. The intensity $[(p_2'\ 2\ 1|S^{\frac{1}{2}}|s_4\ 1\ 0)]^2$ could then be easily found by substituting 0.395 for 0.919 and 0.919 for -0.395; and the intensity

$$[(p_4'\ 2\ 1|S^{\frac{1}{2}}|s_2\ 1\ 0)]^2$$

could be found by substituting 0.266 for 0.964 and -0.964 for 0.266.

These calculations have been carried out in detail for all of the members of the multiplet $2p^54p\rightarrow2p^53s$ which show the Paschen-Back effect. The results are shown in Fig. 1. They involve transitions between p_4, p_2 and all the s_i-levels. In Fig. 1: (A) represents λ3370 $3s_3-4p_{2,4}$ ($j=2\rightarrow j=1, 2$); (B) represents λ3418 $3s_4-4p_{2,4}$ ($j=1\rightarrow j=1, 2$); (C) represents λ3460 $3s_3-4p_{2,4}$ ($j=0\rightarrow j=1, 2$), the forbidden triplet shows very strongly; (D) represents λ3593 $3s_2-4p_{2,4}$ ($j=1\rightarrow j=1, 2$); (E) represent λ3593 in perpendicular polarization showing clearly the presence of two components almost hidden in (D) by the parallel components. In Figs. 1 and 2 the dots indicate parallel components.

The theoretical patterns are plotted with intensities proportional to the lengths of the lines. At first it was thought that a better comparison would be obtained if the length represented log intensity, but the fact that the grating does not respond equally for parallel and perpendicular light seemed to outweigh this advantage. In comparing theoretical and observed patterns, this fact must be borne in mind, together with a few other considerations. The peak position of a line does not always represent its true position. The microphotometer tends to make a weak line be "attracted" by a strong line close to it. Lines unresolved by the grating tend to add their intensities and give higher peaks than lines which are just on the verge of resolution by the grating which appear simply as broad lines.



FIGS. 1–7. Zeeman patterns in neon. In Figs. 1 and 2 dots represent parallel components; all others are perpendicularly polarized.

FIG.	λ	ORDER	FIELD	FIG.	λ	ORDER	FIELD	FIG.	λ	ORDER	FIELD	FIG.	λ	ORDER	FIELD
1*A*	3370	III	32,000	2*A*	5341–3	II	32,000	4*A*	8300–1	I	32,000	6*A*	8417–8	I	32,000
B	3418	III	32,000	*B*	4708–10	I	32,000	*B*	5748	II	32,000	*B*	8654–5	I	32,000
C	3460	III	32,000	*C*	4425	I	29,000	5*A*	8377–8	I	32,000	7*A*	8136	I	32,000
D	3593	III	32,000	3*A*	8635	I	32,000	*B*	5764	II	32,000	*B*	5656–7	I	29,000
E	3593	II	29,000	*B*	5906	I	29,000					*C*	5719	I	29,000

The Transitions $2p^5md \rightarrow 2p^53p$

In the configurations $2p^5md$ several pairs of levels are close enough to perturb each other very markedly. The asymmetry caused by the interaction of $d_5(j=1)$ and $d_6(j=0)$ has been noted by Murakawa and Iwana[9] for the $2p^53d$ configuration, studying the lines λλ7536 and 7544. They found that the triplet 7544 while remaining symmetrical, shifted toward longer wave-lengths as the field strength increased. We have continued their work with the configurations $2p^54d$, $2p^55d$ and $2p^56d$, ignoring the effect of the level $d_3(j=2)$ which is of second order in comparison with the perturbation between d_5 and d_6, since it is too far from these levels.

Figure 2 illustrates these three members of the series. (*A*) λλ5341–3 are $2p_{10}-4d_{5,6}$; (*B*) λλ4708–10, $2p_{10}-5d_{5,6}$; (*C*) λλ4425, $2p_{10}-6d_{5,6}$. They illustrate very effectively the most important feature of the Paschen-Back effect. The coupling is almost the same for all three of the d_5, d_6 terms so that progression in the series using constant field strength would have the same effect as increasing the field strength on a single line. An inspection of the figure shows the crowding together of the lines in the middle of the pattern and the increasing separation of the lines on the edges of the pattern, together with an enhancement of intensity in the middle of the pattern at the expense of the edges.

The remainder of the figures represent only the perpendicular components. They are shown thus because the presence of the parallel components would obscure some of the results obtained. The parallel components are not shown because under the conditions of excitation it was not possible to obtain these components in sufficient purity, free from contamination by perpendicular components. The reflecting power of the grating being much stronger for polarization parallel to the rulings (perpendicular to the magnetic field) tended to exaggerage this difficulty.

Figure 3 gives two members of the series $2p_7-md_1''(j=1\rightarrow j=2)$ with $m=3$, 4 and show the presence of the forbidden transition $2p_7-md_1'$ $(j=1\rightarrow j=3)$. For $m=3$, the separation between d_1' and d_1'' is 1.77 cm^{-1}, while for $m=4$ it is 1.05 cm^{-1}, consequently the intensity of the "forbidden" components is much greater for $m=4$. Theoretically the intensity of these components should vary as H^4.

Figure 4 represents two members of the series $2p_9-md_1'$, d_1'' $(j=3\rightarrow j=3,\ 2)$. The asymmetry of the pattern is clearly indicated by the resolution of the group on the short wave-length side, and the lack of such resolution on the long wave side.

Figure 5 shows the results for two members of the series $2p_9-md_4$, d_4' $(j=3\rightarrow j=3,\ 4)$. The separations here are 1.83 cm^{-1} for $m=3$, and 1.09 cm^{-1} for $m=4$. Only slight changes in intensity are noted; they should be proportional to H^2.

Figure 6 represents two examples so well resolved that they constitute practically perfect agreement between theory and experiment. Fig. 6(*A*) is λλ8417–8, $2p_8-3d_1'$, d_1'' $(j=2\rightarrow j=3,\ 2)$, and Fig. 6(*B*) is λλ8654–5, $2p_4-3s_1'''$ s_1'''' $(j=2\rightarrow j=3,\ 2)$.

In Fig. 7 is illustrated a very great change in two successive members of a series. Fig. 7(*A*), λ8136, $2p_7-3s_1''''$ $(j=1\rightarrow j=2)$ shows the presence of the forbidden line $2p_7-3s_1'''$ $(j=1\rightarrow j=3)$. In the next member of the series λλ5656–7 the level $4s_1''$ comes very close to $4s_1'''$ and $4s_1''''$ (the total spread of the three levels is 1.97 cm^{-1}, while $3s_1'''-3s_1''''$ $=1.57$ cm^{-1} and $3s_1''-3s_1'''=11.33$ cm^{-1}), and we have the case of three mutually perturbing levels. The very close agreement between theory and experiment for this complicated situation is extremely gratifying. Fig. 7(*C*) is λ5719, $2p_6-4s_1''$, $4s_1'''$, $4s_1''''$ $(j=2\rightarrow j=2,\ 3,\ 2)$.

[9] K. Murakawa and T. Iwana, Tokyo Inst. Phys. Chem. Res. **13**, 283 (1930).

Electron Diffraction Studies of Thin Films

I. Structure of Very Thin Films

L. H. GERMER
Bell Telephone Laboratories, New York, New York

(Received May 8, 1939)

Films of metals and inorganic compounds, deposited upon organic supporting foils by vaporization in high vacuum, have been investigated by electron diffraction by the transmission method. Mean thickness of each of these films has been calculated from the total amount of vaporized material, distance of the film from the source, and calibration of the apparatus based upon weighing the material making up a known area of film. Films of the following substances have been studied: Au, CsI, Sb_2O_3, Pd, Cu, CuCl, Ni, CaF_2, Al, Mg, NaF and Be. Most of these films have been extremely thin, two-thirds of them less than 30A and many less than a single layer of atoms or molecules. The masking effect of the superposed diffraction pattern from the supporting foil is so slight that satisfactory patterns can be obtained from films of mean thickness as small as one layer of atoms if their atomic number is as large or larger than that of copper (29); in the case of beryllium (4) films as thick as 50A are required. Diffraction patterns obtained are characteristic of three-dimensional crystals. The ionic compounds studied produce films made up of rather large crystals, 100A or more on a side, even in the very thinnest films. Metal films are, in general, made up of smaller crystals than films of ionic compounds. *The observations prove that atoms and molecules are able to move over the surface of the supporting foil, and that they are drawn together into crystals of considerable size.* Crystals in metal films have been randomly oriented, except for slight orientation in aluminum films; crystals of some ionic compounds have been found to possess a strong preference for some particular orientation, and others have been randomly oriented. No correlation has been observed between average shape of crystal and degree of orientation. Intensity anomalies have been observed in diffraction patterns from calcium fluoride and cuprous chloride. Applications of the method of this investigation to study of corrosion and chemical change are suggested.

ELECTRON diffraction patterns obtained by the transmission method are often more satisfactory than those obtained by reflection. This paper is devoted to the description of a procedure for preparing films for investigation by this method, to studies of some representative films, and to an indication of fields of usefulness of such studies.[1]

The experiments described here were undertaken for the purpose of determining the structure of extremely thin films, and of discovering the limit of sensitivity of the method. Diffraction patterns have been obtained from films of many different substances and many different thicknesses; each film has been deposited by vaporization upon a supporting foil of organic material. The lower limit upon the mean film thickness which can be investigated successfully is determined by the structure of the film and the scattering powers of its atoms, and by the thickness and composition of the supporting foil which gives a diffraction pattern of its own superposed upon that of the film being studied. The amorphous supporting foils which are used are so thin that many satisfactory diffraction patterns have been obtained from films of average thickness less than a single layer of atoms or molecules. In most of these cases and in many others the average dimensions of crystals, both in the plane of the supporting foil and normal to it, are much larger than the mean film thickness.

PREPARATION OF FILMS

Each film was prepared by vaporization in high vacuum from a V-shaped tungsten ribbon upon a thin foil of organic material supported across a narrow slit. An enlarged photograph of this slit[2] is reproduced as Fig. 1.

[1] Many experimental papers describing electron diffraction experiments by the transmission method have been published since the pioneer work of G. P. Thomson. The most notable of the comprehensive investigations seem to be those of F. Kirchner; Physik. Zeits. **30**, 1025–1028 (1930); Zeits. f. Physik **76**, 576–596 (1932).

[2] The material of this slit is a copper-beryllium alloy. Slits used in earlier experiments were made of a nickel alloy. With this earlier alloy it was discovered that the organic foil spread across the slit often had upon its surface

The supporting foil is prepared from a chloroform solution of the plastic known as Formvar No. 1595, manufactured by the Shawinigan Products Corporation.[3] A drop of this solution is first placed upon a slightly greasy microscope slide, then immediately spread over the surface of the slide by means of a razor blade. By maintaining a predetermined clearance between blade and surface of the slide the thickness of the resulting foil of Formvar can be made fairly uniform and have any desired value. Most of the foils have been about 200A thick although 100A foils have occasionally been used. These values of thickness were determined by optical measurement of representative foils (upon chromium) using polarized sodium light in the method devised by Blodgett.[4]

After a Formvar foil has been deposited upon a glass slide, the slide is lowered at an inclination of about 30° to the horizontal into a beaker of water. The Formvar foil floats off on the water, and is captured from the water surface across the slit of Fig. 1.

Vaporization upon this foil of the material to be studied is carried out in high vacuum by a convenient technique devised by Dr. W. H. Brattain of these Laboratories. A tungsten ribbon 4 mm wide and 125 mm long is bent lengthwise into a narrow V. This is cleaned by a preliminary heating in vacuum for 10 seconds at a current of 52 amp., corresponding to a temperature of slightly over 3000°K. The slit with its Formvar foil is supported horizontally at an appropriate distance, 14 to 35 cm, above this V-shaped ribbon, and material placed in the V of the ribbon is vaporized upon it by heating for 5 seconds at 40 amp., corresponding to about 2600°K. (Palladium was vaporized at 50 amp.)

Mean thickness t of a vaporized film prepared in this manner is calculated from the amount of material M, its density δ, and the distance l of the Formvar foil above the V-shaped ribbon, by means of the formula $t=kM/l^2\delta$. The value of the constant k was found to be 0.45 by direct calibration of the apparatus; this consisted of weighing the film deposited upon a horizontally placed glass slide of known area by evaporation of a known amount of material. The data of this calibration are written down in Table I. It is interesting to point out that the constant $k=0.45$ is about 50 percent greater than the value, $1/\pi$, which it would have if the Formvar foil were mounted parallel to a surface from which evaporation from a point source satisfied the cosine law.

TABLE I. *Calibration of vaporization apparatus.*

Vaporized Material Mass M	Deposited Film Area a	Dist. l	Mass m	$k=ml^2/Ma$
Gold 59.6 mg	6.4 cm²	7 cm	3.39 mg	0.435
Gold 59.6	6.4	14	0.89	0.457
Cs I 70.8	6.4	17.5	0.674	0.455

Film Thickness and Mean Crystal Size

Diffraction patterns have been obtained from more than 100 different films of 12 different substances. In Table II are given data relating to about half of these films; the elements or compounds are arranged from gold to beryllium in order of decreasing atomic numbers. For convenience of reference the various films are designated by numbers in column 2. Diffraction patterns from some of the films are reproduced in this paper; the figure numbers of these are given in column 3. In column 7 is the mean thickness of each film as calculated from the data of columns 4, 5 and 6 by means of the formula $t=0.45\times10^{-8}M/l^2\delta A$. For many of the films mean crystal sizes have been estimated from microphotometer curves of diffraction rings. Widths of rings at half-maximum are written down in column 9, and in column 11 corresponding estimates of crystal dimensions. Finally, in the last column are given the "*Strukturbericht*" designations of the types of crystal structure corresponding to diffraction patterns obtained from the various films.

Films of gold

Gold films of mean thickness 22A, 10A and 5.2A, described on the first three lines of Table II, gave diffraction patterns characteristic

oriented crystals of $Ni(OH)_2$, which produced diffraction patterns that were very annoying. [L. H. Germer, Zeits. f. Krist. **100**, 277–284 (1938)].

[3] In a few cases foils of Resoglaz have been used. These are prepared from a benzene solution of Resoglaz in the same way Formvar foils are prepared from a chloroform solution of Formvar.

[4] K. B. Blodgett, J. Phys. Chem. **41**, 975–984 (1937).

TABLE II. *Films formed by vaporization.*

1	2	3	4	5	6	7	8	9	10	11	12
						Mean Thickness	Microphotometer		Mean Crystal Dimensions		
		Figs.	M (mg)	l (cm)	δ (g/cc)	t (A)	Ring	Width (mm)	(Lower Limit) (A)	(Probable Value) (A)	Structure
Au (79)	1	2	1.88	14	19.3	22	422	0.70	33	50	$A1$
	2		1.88	21		10					$A1$
	3	3, 4	0.44	14		5.2	311	*1.3	—	30	$A1$
	4		0.30	14		3.6			Partly Amorphous		
	5		0.69	22		3.3			Partly Amorphous		
			0.67	22		3.2					
	6		0.44	21		2.3			Partly Amorphous		
	7	5	0.69	31		1.7			Partly Amorphous		
			0.67	31		1.6					
	8		0.30	21		1.6			Doubtful		
	9		0.36	24.5		1.4			Partly Amorphous		
			0.35	24.5		1.4					
	10		0.10	14		1.2			Doubtful		
	11		0.36	28.5		1.0			Partly Amorphous		
			0.35	28.5		1.0					
	12		0.10	21		0.5			No Gold		
CsI (55, 53)	13		7.06	25	4.51	113					$B2$
	14		7.06	35		57					$B2$
	15	6	1.87	25		30	510	0.27	87	200	$B2$
	16		1.87	35		15					$B2$
	17		0.28	19		7.7					$B2$
	18	7	0.28	27		3.8	310	0.38	62	150	$B2$
	19		0.15	24		2.6					$B2$
	20		0.15	29		1.8	211	*0.4	—	100	$B2$
Sb_2O_3 (51, 8)	21		4.0	16	5.67	124					
	22	9	0.4	14		16		0.35	67	150	
	23		0.45	25		5.7					
	24		0.45	35		2.9		*0.4	—	100	
Pd (46)	25		14	22.5	11.7	†53	220	0.66	35	50	$A1$
	26		3.5	22.5		†13	422	0.74	32	50	$A1$
	27		1.80	35		5.6					—
	28		0.70	30		3.0					—

* These values were estimated from visual comparisons with other patterns for which more satisfactory microphotometer curves were obtained.
† The palladium films 25 and 26 were formed upon Formvar foils inclined by 30° to the line of vaporization.

of face-centered cubic crystals, although the pattern from the thinnest of these is so diffuse that one could not identify it with certainty if it were not known to be due to gold. This pattern, and that from the film of 22A thickness, are reproduced as Figs. 3 and 2, respectively. These, and all other photographs of diffraction patterns, including Fig. 4, are reproduced here slightly larger than three-fourths full size.) From relative widths of diffraction rings in these figures it is clear that the mean crystal size was much smaller in the thinner film. (In these patterns the innermost diffuse rings are due to the supporting foil. These rings are more prominent in Fig. 3 than in Fig. 2 because of the longer exposure required to produce the pattern of the thinner gold film.) Diffraction patterns from even more tenuous gold films, numbers 4 to 12, are made up of rings which are still more poorly defined than those of Fig. 3.

In the case of four films, numbers 5, 7, 9 and 11, an artifice was used to strengthen the pattern from the gold relative to that from the Formvar supporting foil; gold was vaporized upon both sides of the Formvar so that the foil formed the center of a sandwich between layers of gold of equal mean thickness. In this way the intensity of the film pattern was sensibly doubled. The improvement is illustrated by the fact that patterns from films number 8 and 10, 1.6A and 1.2A thick, respectively, show only doubtful evidence of the presence of any gold at all and patterns from film number 12, 0.5A thick, are quite indistinguishable from those from Formvar

TABLE II.—*Continued.*

1	2	3	4	5	6	7	8	9	10	11	12
						Mean Thickness	Microphotometer		Mean Crystal Dimensions		
		Figs.	M (mg)	l (cm)	δ (g/cc)	t (A)	Ring	Width (mm)	(Lower Limit) (A)	(Probable Value) (A)	Structure
Cu (29)	29	16	7	16	8.92	138	200	0.56	42	100	$A1$
							220	0.52	45	100	$C3$
	30	17	1.24	14		32					$C3$, $A1$
	31	18	1.24	21		14	220	1.2	19	25	$C3$
	32		0.36	14		9.3					$C3$
	33	19	0.36	21		4.1					$C3$
	34		0.39	29		2.3					$C3$
CuCl	35	10A	‡	17.5	3.53	180	422	0.22	107	250	$B3$
(29, 17)	36	10B	Same after 5 days				—	0.28	84	200	—
							220	0.25	94	250	$B3$
Ni (28)	37		3.38	25	8.90	27	311	1.0	24	30	$A1$
	38		3.38	35		14	—	1.5	15	20	—
CaF_2 (20, 9)	39	11	10	16	3.18	553	620	0.38	64	150	$C1$
	40		1.9	20		67					$C1$
	41		1.9	30		30	400	0.46	51	100	$C1$
Al (13)	42	20	1.27	14	2.70	108	422	0.40	59	100	$A1$
	43		1.27	21		48					$A1$
	44		0.33	14		28	220	0.44	53	100	$A1$
	45		0.33	21		12	Amorphous				—
Mg (12)	46		—	—	1.74	500	$21\bar{3}3$	0.35	67	150	$A3$
Na F	47	12	18	18	2.79	896	420	0.33	71	150	$B1$
(11, 9)	48	13A, B	3.0	20		121	420	0.33	71	150	$B1$
	49		3.0	30		54	420	0.34	69	150	$B1$
	50	14A, B	0.6	20		25	400	0.34	69	150	$B1$
	51	15	0.6	30		11					$B1$
	52		0.2	20		8	220	*0.5	39	100	$B1$
Be (4)	53		1.00	14	1.823	126	$10\bar{1}1$	0.55	43	100	$A3$
	54		1.00	21		56	$10\bar{1}1$	0.53	44	100	$A3$

* This value was obtained from visual comparison with other patterns for which more satisfactory microphotometer curves were obtained.
‡ The film of CuCl, number 35, was prepared by vaporization of 7.5 mg of $CuCl_2 \cdot 2H_2O$, which decomposes on heating to produce CuCl.

alone; yet patterns from the double films 7, 9 and 11, of comparable mean thickness, are fairly clear.

Patterns from film number 5, 3.3A of gold on one side of the Formvar foil and 3.2A on the other side, resemble Fig. 3 rather closely. There is, however, a difference; on patterns from film number 5 the (111) Debye-Scherrer ring is relatively much too strong. This is more marked on the pattern from film 7, reproduced as Fig. 5, and still more marked on patterns from films 9 and 11. On the latter patterns no trace can be found of the (200) Debye-Scherrer ring, although there are clearly defined diffuse rings corresponding to the Miller indices (311) and (331)/(420) and weak indications of two other rings corresponding to (422)/(333)/(511) and to (531)/(600)/(442).

The films which produced patterns on which the ring at the (111) position is exceptionally strong are described in Table II as "partly amorphous." These films certainly contained some exceedingly small crystals, but they also contained gold atoms which produced a diffraction pattern consisting chiefly of a single diffuse ring corresponding by the Bragg formula to a spacing of 2.3A [the (111) ring of gold]. Although the mean size of the exceedingly small crystals has not been estimated, it was certainly very much less than 25A; this conclusion is based on the fact that rings of a face-centered cubic structure are fairly clear and prominent

FIG. 2. Diffraction pattern from gold film of 22A mean thickness (number 1, Table II).

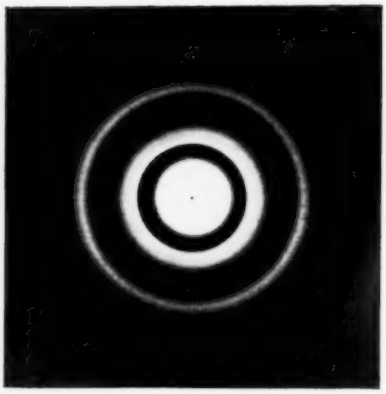


FIG. 3. Pattern from gold film of 5.2A mean thickness (number 3).

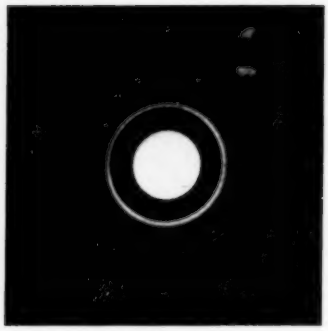


FIG. 5. Pattern from gold film number 7 (1.7A of gold on one side of Formvar foil and 1.6A of gold on the other side).

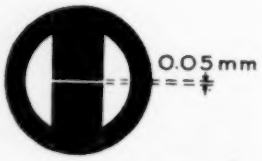


FIG. 1. Photograph of slit used as support.

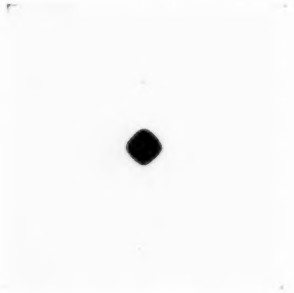


FIG. 4. Reproduction of the center of Fig. 3. enlarged 14.5 times, showing the shape of the primary electron beam.

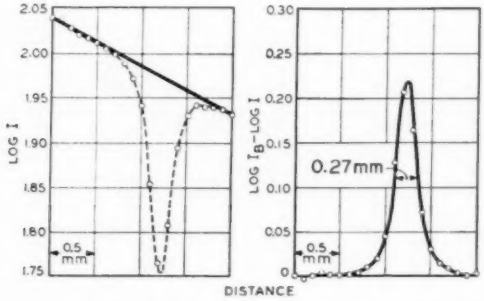


FIG. 8. Microphotometer curves of the (510) diffraction ring on the pattern from the caesium iodide film of mean thickness 30A (Fig. 6).

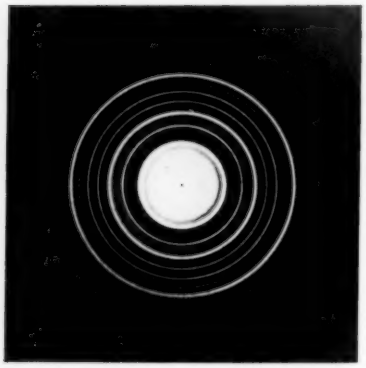


FIG. 6. Pattern at normal incidence from caesium iodide film of 30A mean thickness (number 15).

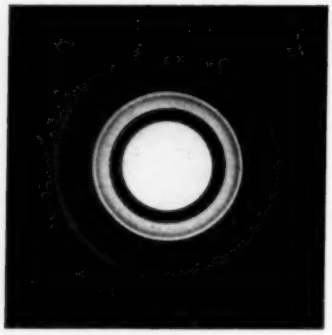


FIG. 7. Pattern at normal incidence from caesium iodide film of 3.8A mean thickness (number 18).

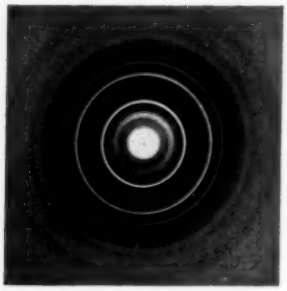
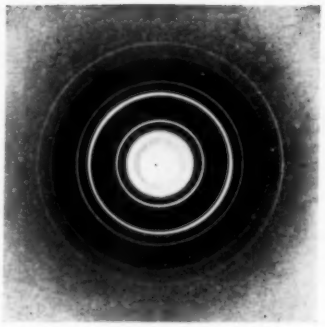
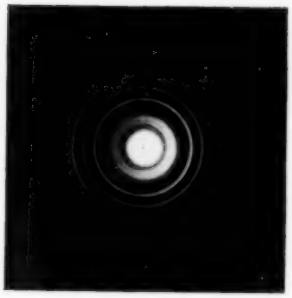


FIG. 9. Pattern at normal incidence from antimony oxide film of 16A mean thickness (number 22).



A



B

FIG. 10. Patterns at normal incidence from cuprous chloride film of 180A mean thickness. A. Freshly prepared (number 35). B. After five days (number 36).

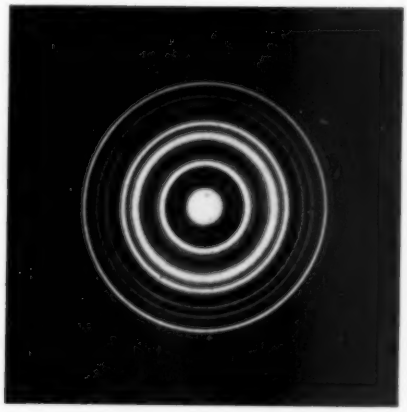


FIG. 11. Pattern at normal incidence from calcium fluoride film of 553A mean thickness (number 39).

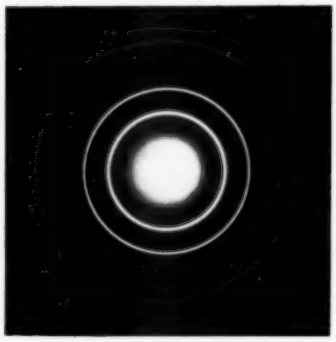
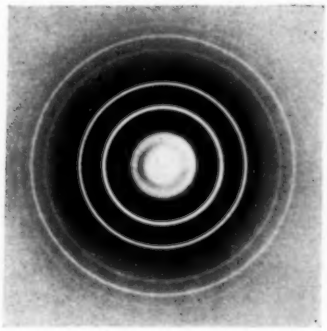
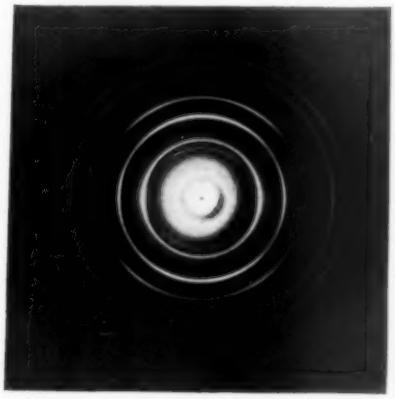


FIG. 12. Pattern at normal incidence from sodium fluoride film of 896A mean thickness (number 47).

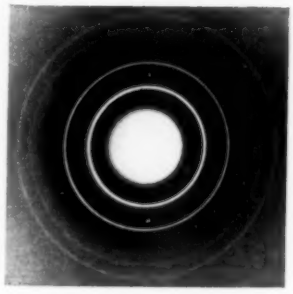


A

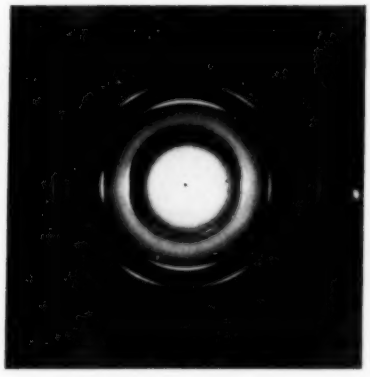


B

FIG. 13. Patterns from sodium fluoride film of 121A mean thickness (number 48). A. Normal incidence. B. 45° incidence (rotated about horizontal axis).



A



B

FIG. 14. Patterns at normal incidence and at 45° incidence from sodium fluoride film of 25A mean thickness (number 50).

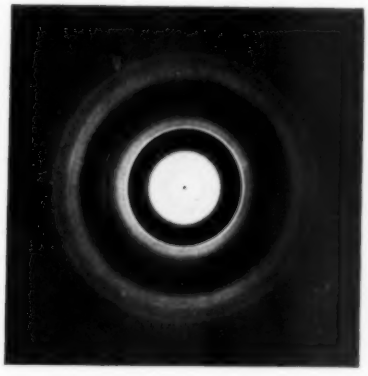


FIG. 15. Pattern at normal incidence from sodium fluoride film of 11A mean thickness (number 51).

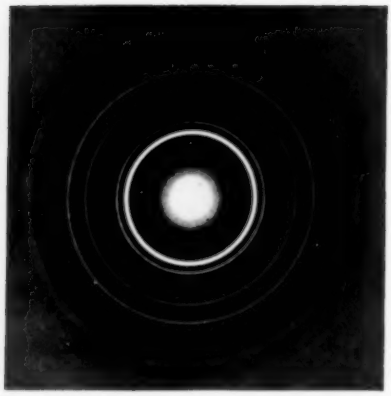


FIG. 16. Pattern from freshly prepared copper film of 138Å mean thickness (number 29), showing the presence of some cuprous oxide.

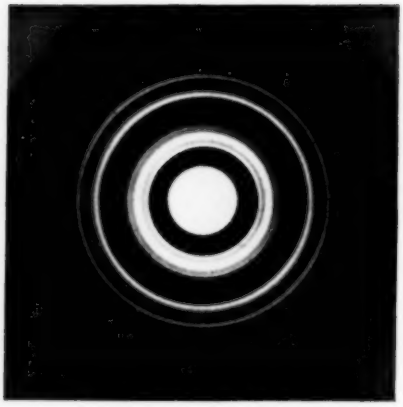


FIG. 17. Pattern from freshly prepared copper film of 32Å mean thickness (number 30), showing rings due to oxide and rings due to copper of about equal strengths.

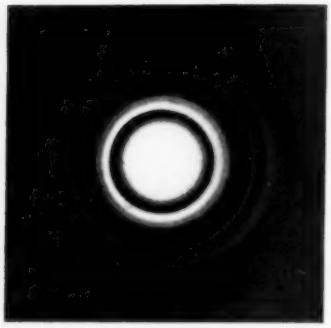


FIG. 18. Pattern from freshly prepared copper film of 14Å calculated mean thickness (number 31), showing only rings due to oxide.

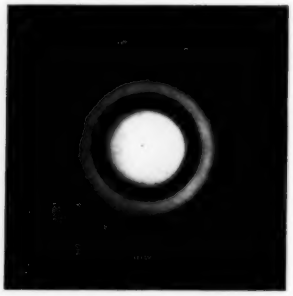


FIG. 19. Pattern from copper film of 4.1Å mean thickness (number 33). Rings due to oxide can still be identified.

on a pattern from a film of cuprous oxide only 2.3A thick in which the mean crystal size was less than 25A (film number 34).

The average size of the crystals in directions normal to the primary beam is calculated from broadening of diffraction rings by means of the Scherrer formula. For fast electrons this formula can be written in the approximate form $C=L\lambda/\Delta R$, where C is the mean crystal size normal to the primary beam, $L\lambda=2.3\times10^{-6}$ mm² is the product of the specimen-plate distance and the electron wave-length, and ΔR is that part of the width of a diffraction ring at half-maximum attributable to finite size of crystals.

To determine mean crystal size one must measure widths of diffraction rings and must know the size of the primary electron beam at the photographic plate. The latter can be photographed directly, but this is often unnecessary; on a diffraction pattern obtained from a very thin film, which requires a relatively long exposure, the spot where the primary beam strikes the plate is reversed and the shape of the beam can sometimes be seen clearly at the center of the pattern. This spot appears at the center of Fig. 3, and in Fig. 4 the same spot is reproduced again after an enlargement of 14.5 times.

It is interesting to compare the spot of Fig. 4 with that to be expected from the electron optical system which produced the primary beam. Electrons from a tungsten filament were accelerated up to a square pinhole, which has vertical and horizontal edges 0.1 mm long, and beyond this they traveled for 84 cm in a space free from electric fields passing through a

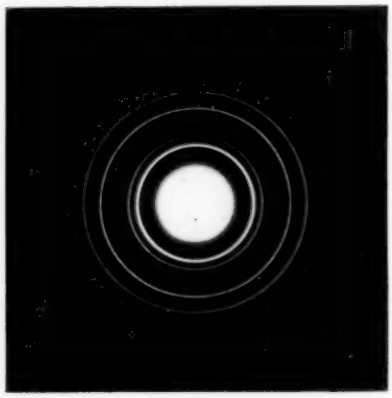


FIG. 20. Pattern from aluminum film of 108A mean thickness (number 42).

magnetic focusing coil located at a point one fourth the distance from pinhole to photographic plate.[5] If the optical system had been perfect the image of the pinhole on the plate would have been enlarged by a factor of approximately three and would have been somewhat rotated; thus the spot of Fig. 4 would have been a square with a side of $0.1\times3\times14.5=4.3$ mm, and diagonal of $4.3\sqrt{2}=6.1$ mm. Measurement of the actual spot gives a diagonal of 6 mm, with the sides of the square bulged outward to make it somewhat barrel-shaped. This deformation was undoubtedly due to imperfection of the electron optical system.

The primary beam spot on the photographic plate is usually much smaller than that of Figs. 3 and 4, and is not approximately square. This would not be possible if the primary pinhole were a "diffuse" source of electrons. If, however, electrons came through the pinhole along parallel lines, as they do approximately, the hole would be analogous to an aperture illuminated by an infinitely distant source of light. It would then be possible for the magnetic electron lens to bring the electrons to a point focus. It is apparent from a number of observations of the primary electron beam that this condition was usually more nearly realized than was the formation of a true image of the pinhole. When the spot was extremely small it was always found to have a badly distorted shape which must have been due primarily to the fact that some electrons coming through the pinhole were moving in directions inclined to the line of motion of most of them.

Although size and shape of the primary beam were often noted no systematic observations were carried out. Thus estimates of mean crystal size from widths of diffraction rings are subject to this added uncertainty. In column 10 of Table II are given values of mean crystal size calculated directly by the Scherrer formula from widths of rings at half-maximum (column 9), and in column 11 values of crystal size calculated after the width at half-maximum has been reduced by the amount which is crudely guessed to be the size of the primary beam.

[5] This is the diffraction camera which was described earlier [L. H. Germer, Rev. Sci. Inst. **6**, 138–142 (1935)], modified by the addition of the magnetic focusing arrangement and by some minor changes.

The gold films 1 and 3 are estimated in this manner to have been made up of crystals of mean size 50A and 30A, respectively. These crystals were randomly oriented, and the mean dimension normal to the film surface was in each case about the same as parallel to it. These conclusions were reached from comparisons of patterns obtained with the films inclined by 45° to the primary beam direction and normal to it; it was found that these patterns are identical.

Films of some inorganic compounds

Many films of inorganic compounds that have been examined have produced diffraction patterns made up of rings which are sharper than those from films of gold. Although this is not true of all compounds studied (for example, silicon dioxide, lead oxide and barium borate), it is true of those written down in Table II. Rings of the patterns of Figs. 6 and 7 from caesium iodide films are, for example, much sharper than those of any of the patterns from gold. The mean thicknesses of the films that produced these patterns were, respectively, 30A and 3.8A. A diffraction pattern was obtained from each film with the primary beam inclined by 45° to the surface normal as well as parallel to it. The patterns obtained at 45° are made up of rings along which intensities vary somewhat, indicating slight preferential orientation among the caesium iodide crystals. These rings are, however, as far as one can judge, of perfectly uniform width. This uniformity proves that the mean crystal size was in each case about the same in the plane of the film and in directions inclined to it by 45°. Thus in both of these films, and in all the other caesium iodide films as well, the crystals had about the same mean dimension in all directions.

Microphotometer curves were obtained from the (422), (510), (521), (440), (530) and (600)/(442) rings on the pattern of Fig. 6, and from the (211), (220) and (310) rings on the pattern of Fig. 7. One of these curves, from the (510) ring of the pattern of Fig. 6, is reproduced as Fig. 8. On the left are plotted as ordinates logarithms of the light intensity transmitted through the photographic plate, and at the right differences between these values and the background curve which is drawn in on the curve at

the left. The width at half-maximum is found to be 0.27 mm which agrees fairly well with values obtained from microphotometer curves of other rings of this pattern. The microphotometer curve of the (310) ring of the pattern of Fig. 7 has a width at half-maximum of 0.38 mm. These values, and ring widths of patterns from other films, are written down in column 9 of Table II, and the ring upon which each of the measurements was made is designated in column 8 by its Miller indices. In most cases the values obtained were checked by measurements upon other rings as well. In measurements upon very thin films it was necessary in each case to choose rings for which the diffuse rings due to Formvar do not introduce large errors into measured widths. Representative diffraction patterns were chosen for measurement; those which were not measured have rings of comparable widths.

The diffraction rings of Figs. 6 and 7, as well as those of patterns from other caesium iodide films, are somewhat narrower at half-maximum than the total width of the primary beam of Fig. 3. It seems clear that the beam must have been narrower when these patterns were obtained, and the estimated mean crystal dimensions of 200A and 150A given in column 11 take this into account. The diffraction pattern from the thinnest caesium iodide film, number 20 of mean thickness 1.8A, is not reproduced here because the diffraction rings from the Formvar supporting foil are relatively so very prominent that the rings from caesium iodide are unsuitable for exhibition purposes. These rings are, however, very sharp and correspond to crystals about 100A on a side both in the plane of the film and normal to it.

In Fig. 9 is reproduced a diffraction pattern obtained at normal incidence from a film of antimony oxide, film number 22 of mean thickness 16A. The rings of this pattern are extremely narrow and correspond to mean crystal dimensions of about 150A in the plane of the film. A pattern from the same film obtained with the primary beam inclined by 45° to the surface normal differs from that of Fig. 9 in two respects; the diffraction rings are no longer of uniform intensities around their circumferences, indicating a slight degree of preferential orientation among the crystals, and the widths also are not uniform. Along the diameter of the pattern that is parallel to the axis about which the film was rotated, the diffraction rings are as narrow as in Fig. 9, but along the diameter of the pattern at right angles to this the rings are definitely wider, although still rather sharp. From comparisons with other patterns it appears that these greater widths correspond to a mean crystal dimension of somewhat less than 100A in directions at 45° to the surface normal. Presumably the mean dimension normal to the film was still smaller by the factor, cos 45°, making this dimension equal to 50A or a little more. Thus the average dimensions of antimony oxide crystals in film number 22 were 150A in the plane of the film, and about 50A normal to it. Other antimony oxide films produced diffraction patterns which were not greatly different, indicating that crystals in all of them had slight preferential orientations and were, on the average, considerably thinner normal to the film surfaces than parallel to them.

Patterns from films of other compounds, cuprous chloride, calcium fluoride and sodium fluoride, appear in Figs. 10–15. These films and others formed from the same compounds are found to have been made up of large crystals, of the order of 100A or 200A on a side, even in those cases in which the average film thickness was very much less than 100A. Each of the cuprous chloride and sodium fluoride films was composed of crystals having nearly the same mean dimension normal to the film and parallel to it, although crystals are randomly oriented in one case and sharply oriented in the other. Some calcium fluoride films, on the other hand, produced diffraction rings at 45° incidence which, although apparently uniform in intensity, are not uniform in width. The patterns indicate that although these crystals were randomly oriented they were on the average lamellar in shape with the small dimension normal to the film.

The patterns from the thinnest films of caesium iodide and sodium fluoride are not exhibited here because these are less intense than the superposed patterns from the Formvar foils and would not show up well in reproduction. Because of the greater scattering power of caesium fluoride relative to sodium fluoride,

patterns from thinner films of the former can be reproduced. Thus one sees that the rings produced by the caesium iodide film 3.8A thick (Fig. 7) have about the same intensity relative to the Formvar pattern as have the rings from the 11A film of sodium fluoride (Fig. 15). Assuming that the Formvar foils were of equal thickness, as they were approximately, one concludes that the caesium iodide and sodium fluoride films had equal scattering powers. This experimental observation can be checked. Consider rings of about the same diameters on the two patterns, corresponding to the same multiplicity of crystal planes, and for each of which the two sorts of atoms of its structure scatter in phase. If the crystal size in each film is assumed uniform the intensities of such rings will be proportional to $I=n^2Z^2/N$, where n is the number of atoms in a film, Z their atomic number and N the number of crystals in the film. For the caesium iodide film, $n=8.0\times10^{14}$ atoms per cm^2, $Z=54$ (the mean of 55 and 53) and N is proportional to the thickness, 3.8A; for the sodium fluoride film, $n=8.8\times10^{15}$ atoms per cm^2, $Z=10$ (the mean of 11 and 9) and N is proportional to the thickness, 11A (assuming that the crystal size is the same in the two films, as it is approximately). The ratio of intensities of the two rings is then

$$I_{CsI}/I_{NaF}=(8.0\times10^{14})^2(54)^2/3.8 \div(8.8\times10^{15})^2(10)^2/11=0.7,$$

which agrees well enough with the crudely estimated ratio of unity.

Films of metals

Many metal films have given diffraction patterns made up of rings much broader than those from any of the ionic compounds in Table II, and some very thin metal films have produced patterns which have not been interpreted.

The two thickest palladium films (numbers 25 and 26) produced rings characteristic of a face-centered cubic structure with ring widths about the same as those from gold films of comparable mean thickness. (See Fig. 2.) Patterns from the two thinnest palladium films (numbers 27 and 28) are made up of rather broad rings which have not yet been interpreted.

Interesting diffraction patterns were obtained from freshly formed copper films of various thicknesses. The pattern of Fig. 16, from film number 29 of 138A mean thickness, corresponds to two different sets of randomly oriented cubic crystals with spacings and intensities appropriate, respectively, to copper and to cuprous oxide. From relative intensities one estimates that about 15 percent of the film was oxide and the rest metallic copper. On the pattern from film number 30 of mean thickness 32A, Fig. 17, the rings due to copper and those due to oxide are of about equal intensity. Films number 31 and 33 of mean thickness, respectively, 14A and 4.1A gave diffraction patterns, Figs. 18 and 19, corresponding to cuprous oxide only. In the thicker of these films the average crystal size was about 25A, and in the thinner film considerably smaller. Rings corresponding to the structure of cuprous oxide can be seen clearly even in the pattern from film number 34 which was only 2.3A thick, although the pattern from the Formvar foil is in this case too prominent for good reproduction.

The thicker nickel film (number 37) produced diffraction rings agreeing with the face-centered cubic structure of metallic nickel, but the pattern from the thinner film (number 38) is not yet understood. Rings on both of these patterns are very broad and the nickel crystals in the thicker film were only slightly larger than those in the gold film which produced the pattern of Fig. 3.

Aluminum films gave diffraction rings, attributable to face-centered metallic aluminum, which are sharper than those produced by gold, palladium, copper or nickel. Rings from an aluminum film only 28A thick, number 44, seem to be just as sharp as those from a film 108A thick, which are reproduced in Fig. 20. A thinner film of aluminum only 12A thick (number 45) seemed to be entirely amorphous, due perhaps to the formation of amorphous Al_2O_3.

Excellent diffraction patterns were obtained from films of magnesium and beryllium. These patterns correspond to large randomly oriented crystals which have the hexagonal structures of the bulk metals. With even the thinnest Formvar foils, which can be handled at present, satisfactory patterns cannot be obtained from beryl-

TABLE III. *Orientation and crystal shape in various films.*

	Degree of Orientation	Variation of Degree of Orientation with Thickness	Crystal Shape
CsI	Slight	Same for all thicknesses	Equiaxed
Sb_2O_3	Slight	Same for all thicknesses	Lamellar
Cu_2O (formed from unoriented Cu crystals)	None	———	Equiaxed
CuCl	None	———	Equiaxed
Cu compound (formed from unoriented CuCl crystals)	Very marked	?	?
CaF_2	None	———	Sometimes Lamellar
NaF	Marked	Varying with thickness	Equiaxed

lium films much thinner than 50A because of the low scattering power of beryllium atoms.

Crystal Orientation

In some aluminum films there was a slight preference for orientation of crystals with a (111) plane parallel to the film. Except for this, no preferential orientation of crystals was found in any of the metal films studied. In films of inorganic compounds, crystals were entirely unoriented in some cases and sharply oriented in others.

Crystals were unoriented in films of cuprous oxide (numbers 31 to 34), cuprous chloride (number 35) and calcium fluoride (numbers 39 to 41). There was a slight preference for a particular orientation in caesium iodide films (numbers 13 to 18) and in antimony oxide films (numbers 21 to 23), and for both substances the degree of orientation did not vary perceptibly with thickness. Diffraction patterns from sodium fluoride indicate that crystals in many of these films were markedly oriented (with a (100) plane tending to lie parallel to the film surface). Orientation which seemed to be still sharper was discovered in an unidentified compound which was formed when cuprous chloride stood in air for several days. (Fig. 10B from film number 36 exhibits many diffraction rings of this compound in a pattern at normal incidence.)

Crystals in the thickest sodium fluoride film (number 47, 896A thick) were unoriented, as they were also in the thinnest films (numbers 51 and 52); the sharpest orientation occurred at a thickness of 25A, Figs. 14A and 14B. The variation of degree of orientation with thickness can be followed by observing relative intensities of various diffraction rings in patterns obtained at normal incidence; the (222) and (422) rings have normal intensities in Figs. 12 and 15, they are somewhat weaker in Fig. 13A, still weaker in a pattern at normal incidence from film number 49, and entirely missing in Fig. 14A. The degree of orientation is, however, more strikingly exhibited by patterns at 45° incidence, such as the patterns of Figs. 13B and 14B. A pattern at 45° incidence from the film of intermediate thickness, number 49, exhibits orientation of an intermediate degree of sharpness.

One might think that marked orientation would occur for crystals of lamellar shape, thinner normal to the film surface than parallel to it, and that randomly oriented crystals would tend to have the same mean dimension in different directions. From the collected data of Table III it appears, however, that there is no such correlation. Some lamellar crystals, thin normal to the film, were unoriented (CaF_2), and some crystals which had the same mean dimension in different directions, designated as "equiaxed," were oriented (CsI and NaF).

Intensity Anomalies

No abnormalities have been observed in *positions* of rings on any of the diffraction patterns which have been interpreted.[6] There is,

[6] It has been stated that in some films made up of extremely small crystals the size of the unit cell does not have the normal x-ray value [G. I. Finch and S. Fordham, Proc. Phys. Soc. **48**, 85–94 (1936); E. Pickup, Nature **137**, 1072 (1936)]. Although I have observed no such anomaly, many of my measurements have not been of sufficient precision to detect the small variations which have been reported.

Quarrell has stated that films produced by vaporization of metals, which normally form face-centered cubic crystals, are sometimes rhombohedral and sometimes hexagonal

however, a curious and unexplained anomaly in the *intensity* of a diffraction ring in patterns from films of cuprous chloride and similar anomalies have been found for a calcium fluoride film.

In patterns from other materials no anomalies are apparent. The diffraction rings obtained from caesium iodide at normal incidence seem, for example, to have normal relative intensities as was to be expected because of the slight degree of orientation of crystals in these films. This is illustrated by data relating to the pattern from caesium iodide film number 15, which are collected in Table IV.

In the first two columns of Table IV are written down measured diameters and estimated intensities of the 11 smallest rings on the pattern of Fig. 6. In the next three columns are given

TABLE IV. *Measured and calculated diameters and intensities for* CsI.

Fig. 6						Calculated	
Measured Diameters (mm)	Estimated Intensities (arbitrary scale)	*hkl*	*H*	*p*	*S*	Diameters (mm)	Intensities (arbitrary scale)
—	—	100	1	3	0.4		0.01
14.7	10	110	2	6	10.3	14.5	19.
—	—	111	3	4	0.3		0.01
20.5	5	200	4	3	7.4	20.6	4.9
—	—	210	5	12	0.2		0.01
25.3	9	211	6	12	5.7	25.2	11.7
29.0	5	220	8	6	4.8	29.1	4.1
—	—	300	9	3	0.07		0.0004
32.5	6	310	10	12	4.2	32.5	6.4
—	—	311	11	12	0.05		0.0009
35.6	3	222	12	4	3.9	35.6	1.8
—	—	320	13	12	0.03		0.0003
38.5	8	321	14	24	3.4	38.5	8.3
41.1	1	400	16	3	3.2	41.1	0.9
—	—	410	17	12	0.04		0.0006
43.7	4	330/411	18	18	3.0	43.6	4.9
—	—	331	19	12	0.04		0.0006
46.0	3	420	20	12	2.7	46.0	2.6
—	—	421	21	24	0.04		0.001
48.2	3	332	22	12	2.5	48.3	2.3

TABLE V. *Estimated and calculated intensities for* CuCl *and for* CaF_2.

hkl	Cuprous Chloride Intensities		Calcium Fluoride Intensities	
	Estimated (Fig. 10A)	Calculated	Estimated (Fig. 11)	Calculated
111	9	6.3	8	7.8
200	—	0.02	*3	0.4
220	10	9.5	10	14.7
311	6	7.0	6	6.4
222	*2	0.03	*2	0.02
400	2	2.3	4	3.8
331	5	4.0	4	3.1
420	—	0.3	—	0.0002
422	5	5.6	6	10.9
511/333	2	3.4	3	2.3
440	1	1.9	1	2.9
531	2	3.8	3	2.3
600/442	—	0.4	—	0.02

Miller indices of the first 20 rings to be expected from crystals of caesium iodide, values of $H=h^2+k^2+l^2$, and values of p the number of cooperating planes. In the sixth column are values of $S=[\mathrm{Cs}]+[\mathrm{I}]\cos\pi(h+k+l)$ characteristic of the caesium iodide structure; here $[\mathrm{Cs}]=(Z-F)_{\mathrm{Cs}}/H$ where Z is the atomic number of caesium and F represents values of the x-ray atomic structure factor taken from the tables of Pauling and Sherman,[7] and $[\mathrm{I}]=(Z-F)_{\mathrm{I}}/H$, the corresponding function for iodine. The calculated diameters are obtained from the relation $D=2L\lambda H^{\frac{1}{2}}/a$, where $a=4.56\times10^{-7}$ mm is the x-ray value for the edge of the unit cube of the caesium iodide structure[8] and $L\lambda=2.345\times10^{-6}$ mm^2 is the independently determined value of the product of the specimen-plate distance and the electron wave-length. Intensities in the last column of Table IV are values of S^2p multiplied by the arbitrary factor 3×10^{-2} so that they can be more conveniently compared with estimated intensities given in the second column. It is clear that positions and relative intensities of rings of the pattern of Fig. 6 agree well with the calculated values. (Nearly equally good agreements are obtained if one omits the F function and writes $[\mathrm{Cs}]=Z_{\mathrm{Cs}}$ in the expression for S and then uses, for the intensity, values of the function S^2p/H. This simple procedure was used in an earlier paper.[2])

in structure—[A. G. Quarrell, Proc. Phys. Soc. **49**, 279–293 (1937)]. This I have not found. One observation regarding the Quarrell paper should be made. In this paper Fig. 4, a diffraction pattern from a film of copper produced by evaporation, is described as due to rhombohedral crystals and the ratios of the spacings corresponding to the second, third, fourth and fifth rings to the spacing corresponding to the smallest are given (Table II) as 0.855, 0.590, 0.499 and 0.476. I have measured diameters on the reproduction and obtained for the first five rings 17.8 mm, 20.7 mm, 29.2 mm, 34.2 mm and 35.8 mm. The ratios of the first of these to the others are 0.860, 0.610, 0.521 and 0.497. Corresponding ratios for face-centered cubic crystals are $(3/4)^{\frac{1}{2}}=0.866$, $(3/8)^{\frac{1}{2}}=0.612$, $(3/11)^{\frac{1}{2}}=0.522$ and $(3/12)^{\frac{1}{2}}=0.500$.

[7] L. Pauling and J. Sherman, Zeits. f. Krist. **81**, 1–29 (1932).

[8] W. P. Davey, Phys. Rev. **21**, 143–161 (1923); G. L. Clark and W. Duane, Phys. Rev. **21**, 380 (1923).

In the case of patterns from cuprous chloride and calcium fluoride measured and calculated values of ring diameters agree equally well, but for some rings measured and calculated intensities do not agree. Data illustrating these disagreements are given in Table V. For cuprous chloride $S^2=X^2+Y^2$, where $X=4[\mathrm{Cu}]\cos\frac{1}{2}\pi(2h+k+l)\cos\frac{1}{2}\pi(h+2k+l)\cos\frac{1}{2}\pi(h+k+2l)+4[\mathrm{Cl}]\cos\frac{1}{2}\pi h\cos\frac{1}{2}\pi k\cos\frac{1}{2}\pi l$, and $Y=-4[\mathrm{Cl}]\sin\frac{1}{2}\pi h\sin\frac{1}{2}\pi k\sin\frac{1}{2}\pi l$; for calcium fluoride $S=4[\mathrm{Ca}]\cos\frac{1}{2}\pi(2h+k+l)\cos\frac{1}{2}\pi(h+2k+l)\cos\frac{1}{2}\pi(h+k+2l)+8[\mathrm{F}]\cos\frac{1}{2}\pi h\cos\frac{1}{2}\pi k\times\cos\frac{1}{2}\pi l$. (In these expressions, as in the corresponding expression for caesium iodide, the symbols in the square brackets represent values of the function $(Z-F)/H$.) The calculated intensities written down in Table V are values of S^2p multiplied by the arbitrary factors, 3×10^{-2} for cuprous chloride, and 5×10^{-2} for calcium fluoride.

It is clear from Table V that the (222) diffraction rings from cuprous chloride and from calcium fluoride are about 100 times as strong as expected, and the (200) ring from calcium fluoride about 10 times as strong. (Asterisks are used in the table to call attention to these extraordinary intensities.) Anomalous strength of (222) reflections has been observed in patterns from a great many different cuprous chloride films, some of which were prepared by the direct reaction of copper and chlorine. Some calcium fluoride films have, however, given rings having more nearly normal intensities; the (440) ring is then stronger and appears normal. No explanation of these phenomena has yet been found.[9]

Some Applications

Crystal growth in thin films

It has been pointed out that in many of the thinner films the dimension of the average crystal in a direction normal to the film was much greater than the mean film thickness. Molecules of caesium iodide and antimony oxide, for example, formed relatively large three-dimensional crystals in films of mean thickness less than that of one molecular layer, and molecules of cuprous chloride, calcium fluoride and sodium fluoride formed large crystals in the thinnest films studied. These observations prove that atoms or molecules, deposited by vaporization, were able to move about on the surface of the Formvar foil and that they actually agglomerated into three-dimensional crystals, in many cases leaving a large fraction of the foil completely bare.

Diffusion of atoms upon surfaces has been deduced from measurements of resistance of extremely thin metal films,[10] from thermionic[11] and photoelectric measurement[12] upon composite surfaces, and from optical observations.[13] It is clear that electron diffraction offers an additional and rather direct method of investigating this type of diffusion.[14] The work reported here was, however, not undertaken with this in mind. Although the deductions which can be drawn from it are naturally extremely fragmentary, it is probably worth while to review some of them.

From Table II one sees that the ionic compounds studied (excepting only cuprous oxide produced from copper) form considerably larger crystals than the metals, and if one looks up melting points it appears that there is a tendency for low melting point substances to form larger crystals than those having high melting points. In films of ionic compounds mean crystal size varies only moderately with film thickness, whereas in gold films crystal size varies greatly with thickness. A test could not be carried out upon extremely thin films of palladium, copper, nickel or aluminum because all of these metals apparently formed chemical compounds, of which only cuprous oxide has been identified. In none of the experiments was there any control of the temperature of a film during vaporization, and

[9] Intensity anomalies in electron diffraction patterns from zinc oxide have been observed and explained by H. J. Yearian, Phys. Rev. **48**, 631–639 (1935). It is not clear, however, that the abnormalities reported here can be interpreted in a similar way.

[10] E. T. S. Appleyard, Proc. Phys. Soc. **49**, Extra Part, 118–135, (Aug. 31, 1937).

[11] W. H. Brattain and J. A. Becker, Phys. Rev. **43**, 428–450 (1933).

[12] R. C. L. Bosworth, Proc. Roy. Soc. **A150**, 58–76 (1935).

[13] E. N. daC. Andrade and J. G. Martindale, Phil. Trans. Roy. Soc. **235**, 69–100 (1935).

[14] J. Krautkrämer [Ann. d. Physik **32**, 537–576 (1938)] has recently used electron diffraction to investigate films of metals deposited upon quartz by vaporization. He has concluded, in agreement with the observations of the present paper, that crystals of appreciable size occur even in the thinnest films.

those films which were only 14 cm from the V-shaped tungsten ribbon must have been much warmer than those at 35 cm. The only quantitative datum is furnished by knowledge of the softening point of Formvar supporting foils, about 250°C; it seems almost certain that this temperature was never exceeded.

The data seem to prove that molecules of the ionic compounds studied were able to move rather freely over Formvar surfaces, and that crystals of these compounds smaller than about 100A on a side were unstable, losing molecules to larger crystals by diffusion. It seems probable that for any compound upon any surface there is in general a critical temperature below which diffusion will not occur,[10] which may, of course, in some cases be above the melting point of the surface. Above this critical temperature there may be a minimum crystal size which is just stable, this size being probably a function of temperature.

The electron diffraction method seems well suited to investigation of the diffusion of molecules over surfaces, and to discovering the relation between temperature and minimum size of crystal. The first experiments might well be the condensation of extremely thin caesium iodide films upon Formvar held at various low temperatures, followed by investigation of crystal size in these films. These experiments might be followed by similar studies of caesium iodide upon beryllium surfaces. Some preliminary and as yet unsatisfactory studies of gold upon beryllium have already been made.

Corrosion and chemical change

Chemical change occurring in thin films can be followed by electron diffraction examination. The chemical and crystalline nature of new compounds formed can often be determined, and quantitative data obtained regarding rate of change.

A study of this sort has been made of a copper film, number 29, 138A thick. From a diffraction pattern produced by the freshly prepared film, Fig. 16, one concludes that about 15 percent of the film was cuprous oxide. After exposure to the air of an ordinary room for 12 days, another pattern was obtained which indicated that the film was then about 50 percent oxide. The oxidation of the copper has now been followed for nearly a year. The amount of oxide varied approximately as the square root of the time, until the film was completely oxidized.

Other studies of copper corrosion under carefully controlled conditions are in progress and under consideration. Corrosion of films of a number of other metals is also being investigated.

Electron diffraction methods can equally well be applied to the examination of chemical changes which cannot be classed as corrosion. In some cases more detailed information can be obtained than one can get in any other way. An example is furnished by the incompleted study of the change of cuprous chloride into some other compound on standing in air. (Films number 35 and 36, Figs. 10A and 10B.) The cuprous chloride crystals are large and completely unoriented. They change into other crystals which are equally large and very sharply oriented. It seems likely that the new compound will be identified in the near future, and that interesting and perhaps valuable information will be obtained regarding the mechanics of the chemical reaction involved.

I am indebted to Mr. W. E. Campbell and Dr. W. H. Brattain for valued advice and assistance. The corrosion studies were undertaken with Mr. Campbell's help and some of them are now being carried out under his direction. The first vaporized films were prepared for me by Dr. Brattain and his method has proved generally useful. I am also indebted to Drs. C. J. Davisson and W. Shockley for many instructive discussions, and to Mr. H. G. Wehe, Mr. K. H. Storks and Mr. A. J. Parsons for assistance with some of the experimental work.

On the Self-Energy and the Electromagnetic Field of the Electron

V. F. WEISSKOPF
University of Rochester, Rochester, New York
(Received April 12, 1939)

The charge distribution, the electromagnetic field and the self-energy of an electron are investigated. It is found that, as a result of Dirac's positron theory, the charge and the magnetic dipole of the electron are extended over a finite region; the contributions of the spin and of the fluctuations of the radiation field to the self-energy are analyzed, and the reasons that the self-energy is only logarithmically infinite in positron theory are given. It is proved that the latter result holds to every approximation in an expansion of the self-energy in powers of e^2/hc. The self-energy of charged particles obeying Bose statistics is found to be quadratically divergent. Some evidence is given that the "critical length" of positron theory is as small as $h/(mc)\cdot\exp(-hc/e^2)$.

I. INTRODUCTION AND DISCUSSIONS OF RESULTS

THE self-energy of the electron is its total energy in free space when isolated from other particles or light quanta. It is given by the expression

$$W=T+(1/8\pi)\int(H^2+E^2)d\mathbf{r}. \qquad (1)$$

Here T is the kinetic energy of the electron; H and E are the magnetic and electric field strengths. In classical electrodynamics the self-energy of an electron of radius a at rest and without spin is given by $W\sim mc^2+e^2/a$ and consists solely of the energy of the rest mass and of its electrostatic field. This expression diverges linearly for an infinitely small radius. If the electron is in motion, other terms appear representing the energy produced by the magnetic field of the moving electron. These terms, of course, can be obtained by a Lorentz transformation of the former expression.

The quantum theory of the electron has put the problem of the self-energy in a critical state. There are three reasons for this:

(a) Quantum kinematics shows that the radius of the electron must be assumed to be zero. It is easily proved that the product of the charge densities at two different points, $\rho(\mathbf{r}-\xi/2)\times\rho(\mathbf{r}+\xi/2)$, is a delta-function $e^2\delta(\xi)$. In other words: if one electron alone is present, the probability of finding a charge density simultaneously at two different points is zero for every finite distance between the points. Thus the energy of the electrostatic field is infinite as

$$W_{st}=\lim_{(a=0)}e^2/a.$$

(b) The quantum theory of the relativistic electron attributes a magnetic moment to the electron, so that an electron at rest is surrounded by a magnetic field. The energy

$$U_{mag}=(1/8\pi)\int H^2d\mathbf{r}$$

of this field is computed in Section III and the result is

$$U_{mag}=e^2h^2/(6\pi m^2c^2a^3).$$

This corresponds to the field energy of a magnetic dipole of the moment $eh/2mc$ which is spread over a volume of the dimensions a. The spin, however, does not only produce a magnetic field, it also gives rise to an alternating electric field. The closer analysis of the Dirac wave equation has shown[1] that the magnetic moment of the spin is produced by an irregular circular fluctuation movement (Zitterbewegung) of the electron which is superimposed to the translatory motion. The instantaneous value of the velocity is always found to be c. It must be expected that this motion will also create an alternating electric field. The existence of this field is demonstrated in Section III by the computation of the expression

$$U_{el}=(1/8\pi)\int E_s^2d\mathbf{r}.$$

There E_s is the solenoidal part (div. $E_s=0$) of the electric field strength created by the electron. The fact that the above expression does not vanish for an electron at rest proves the existence

[1] E. Schroedinger, Berl. Ber. 1930, 418 (1930).

of a solenoidal field[2] apart from the irrotational electric field of the charge. The energies of the electric and magnetic fields of the spin are found to be equal. The spin movement does not, of course, give rise to a radiation. The time average of the Poynting vector is zero.

The electromagnetic field of the spin does not contribute to the self-energy of the electron. It is shown in Section IV, that the charge dependent part of the self-energy to a first approximation is given by

$$W' = \tfrac{1}{2}\int\left(\rho\phi - \frac{1}{c}\mathbf{i}\cdot\mathbf{A}\right)\mathbf{dr} = \frac{1}{8\pi}\int (E^2 - H^2)\mathbf{dr}.$$

Here ρ and $\mathbf{i}$ are the charge and current densities, ϕ and $\mathbf{A}$ are the scalar and vector potential, respectively. If the self-energy is expressed in terms of the field energies, the electric and magnetic parts have opposite signs,[3] so that the contributions of the electric and of the magnetic fields of the spin cancel one another.

(c) The quantum theory of the electromagnetic field postulates the existence of field strength fluctuations in empty space. These give rise to an additional energy, which diverges more strongly than the electrostatic self-energy. The following crude calculation may demonstrate how this particular part of the self-energy arises: Let us consider an electron with radius a. The field fluctuations in a volume a^3 are of the order $E^2 \sim hc/a^4$.[4] The mean frequency of the fluctuations is $\nu \sim c/a$. This field induces the electron to perform vibrations with an amplitude $x \sim eE/m\nu^2$ and an energy $W_{\text{fluct}} \sim e^2E^2/m\nu^2 \sim e^2h/mca^2$. This energy diverges quadratically for infinitely small radius. The exact value is calculated in Section IV and is $W_{\text{fluct}} = \lim_{(a=0)} e^2h/\pi mca^2$.

A new situation is created by Dirac's theory of the positron: The self-energy diverges only logarithmically with infinitely small radius. This fact has been proved[5] only for the first approximation of the self-energy expanded in powers of e^2/hc. However it will be shown in Section VI that the divergence is logarithmic in every approximation. The main purpose of this paper is to show the physical significance of the logarithmic divergence and to demonstrate the reasons of its occurrence.

Let us consider the case of one electron embedded in the vacuum as described by the positron theory. The vacuum is represented by the state in which all negative energy states are filled with electrons. The charge density of these "vacuum electrons" is not observable in the unperturbed state of a field-free vacuum. However, the differences between the actual density and the unperturbed density are observable.

The presence of an electron in the vacuum causes a considerable change in the distribution of the vacuum electrons because of a peculiar effect of the Pauli exclusion principle. According to this principle it is impossible to find two or more electrons in a single cell of a volume h^3 in the phase space. If two electrons of equal spin are brought together to a small distance d, their momentum difference must be at least h/d. This effect is similar to a repulsive force which causes two particles with equal spin not to be found closer together than approximately one de Broglie wave-length.

As a consequence of this we find at the position of the electron a "hole" in the distribution of the vacuum electrons which completely compensates its charge. But we also find around the electron a cloud of higher charge density coming from the displaced electrons, which must be found one wave-length from the original electron. The total effect is a broadening of the charge of the electron over a region of the order h/mc as it is indicated schematically in Fig. 1. The product $\rho(\mathbf{r}-\xi/2)\times\rho(\mathbf{r}+\xi/2)$ is no longer zero for a finite distance ξ, and is given by the function

$$G(\xi) = e^2\frac{mc}{h}\frac{1}{\xi}\frac{\partial}{\partial\xi}\frac{i}{2\pi}H_0^{(1)}(imc\xi/h)$$

(Section II). Here $H_0^{(1)}(x)$ is the Hankel function of first kind. $G(\xi)$ has still a quadratic singularity

[2] A solenoidal electric field is necessarily an alternating field for its time average vanishes in a stationary state, whereas the time average of a magnetic field does not vanish if stationary currents are present.

[3] This at first sight unfamiliar result is connected with the well-known fact that a system of steady currents increases its magnetic field energy if it performs mechanical work, whereas a system of charges decreases its field energy by performing mechanical work.

[4] The fluctuations are of the order of magnitude of the field-strength of one light quantum with wave-length.

[5] V. Weisskopf, Zeits. f. Physik 89, 27 (1934); 90, 817 (1934).

for $\xi=0$. It is shown quantitatively in Section II, that this broadening of the charge distribution is just sufficient to reduce the electrostatic self-energy to a logarithmically divergent expression.

The broadening effect also changes the magnetic field distribution of the spin moment. In positron theory the magnetic field energy is given by

$$U_{\text{mag}}=\lim_{(a=0)}[e^2h/(2\pi mca^2) - e^2mc/(4\pi h)\cdot \lg (h/mca)]. \quad (2)$$

This is equal to the field energy of a momentum distribution spread over a finite region, which is proportional to the spread of charge described above. The divergence, which is less strong than in the one-electron theory,[6] comes from the quadratic singularity of the distribution. The electric field energy of the spin, however, is not equal to the magnetic field energy because of the following effect, which is again based upon the exclusion principle. The vacuum electrons which are found in the neighborhood of the original electron, fluctuate with a phase opposite to the phase of the fluctuations of the original electron. This phase relation, applied to the circular fluctuation of the spin, decreases its total electric field by means of interference, but does not change the magnetic field of the spins since the latter is due to circular currents and is not dependent on the phase of the circular motion. Thus the total solenoidal electric field energy is reduced by interference if an electron is added to the vacuum. The electric field energy U_{el} of an electron in positron theory is therefore *negative* since it is the difference between the field energy of the vacuum plus one electron, and the energy of the vacuum alone. The exact calculations of Section III give $U_{\text{el}}=-U_{\text{mag}}$. Thus the contribution of the spin to the self-energy does not vanish in positron theory and is by Eq. (2)

$$W_{\text{sp}}=-2U_{\text{mag}}=-\lim_{(a=0)}[e^2h/(\pi mca^2) - e^2mc/(2\pi h)\cdot \lg h/(mca)].$$

The broadening effect cannot, however, be applied to the energy W_{fluct}, which is the energy of the action of the electromagnetic field fluctuations upon the electron. The effect of an external field upon an electron in positron theory is to a first approximation the same as one expects for an electron with infinitely small radius, since the effect of the field upon the displaced vacuum electrons can be neglected. For instance, no destructive interference effect would occur in the interaction with a light wave whose wave-length is smaller than h/mc. The exclusion principle does not alter the interaction of an electron with the field as long as one considers that action to a first approximation to be the sum of independent actions at every point; it has only an effect on the probability of finding one particle in the neighborhood of another.

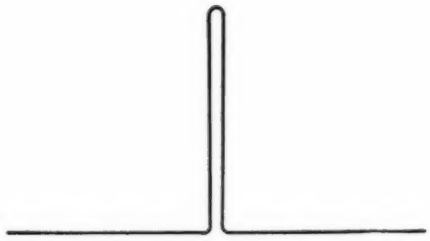


FIG. 1a. Schematic charge distribution of the electron.

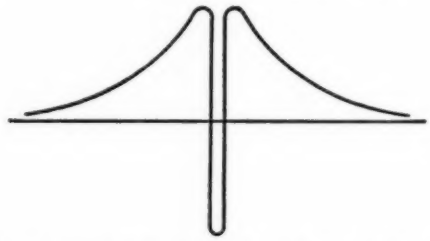


FIG. 1b. Schematic charge distribution of the vacuum electrons in the neighborhood of an electron.

The energy W_{fluct} in positron theory is therefore not different from the same quantity in one electron theory as shown in Section IV. In the former theory, however, it is balanced by the spin energy W_{sp} the most strongly divergent terms of which are just oppositely equal to W_{fluct}. The sum of W_{sp} and W_{fluct} is only logarithmically divergent.

Thus according to positron theory the self-energy of an electron consists of three parts:

(a) The energy W_{st} of the Coulomb field, which diverges logarithmically because of the characteristic spread of charge.

(b) The energy W_{sp} of the oscillatory motion which produces the spin. This energy, although

[6] We use the term "one-electron theory" for the description of the electron by means of the Dirac wave equation without filling up the negative energy states, in order to distinguish it from the "positron theory."

zero in the one-electron theory, is negative and quadratically divergent in the positron theory. This is because of the negative contribution of the magnetic field and the interference effect of the electric field of the vacuum electrons.

(c) The energy W_{fluct} of forced vibrations under the influence of the zero-point fluctuations of the radiation field. The energies (b) and (c) compensate each other to a logarithmic term.

It is interesting to apply similar considerations to the scalar theory of particles obeying the Bose statistics, as has been developed by Pauli and the author.[7] Here the probability of finding two equal particles closer than their wave-lengths is *larger* than at longer distances. The effect on the self-energy is therefore just the opposite. The influence of the particle on the vacuum causes a higher singularity in the charge distribution instead of the hole which balanced the original charge in the previous considerations. It is shown in Section V that this gives rise to a quadratically divergent energy of the Coulomb field of the particle. Thus the situation here is even worse than in the classical theory. The spin term obviously does not appear and the energy W_{fluct} is exactly equal to its value for a Fermi particle.

A few remarks might be added about the possible significance of the logarithmic divergence of the self-energy for the theory of the electron. It is proved in Section VI that every term in the expansion of the self-energy in powers of e^2/hc

$$W=\sum_n W^{(n)} \tag{3}$$

diverges logarithmically with infinitely small electron radius and is approximately given by

$$W^{(n)}\sim z_n mc^2(e^2/hc)^n[\lg\,(h/mca)]^t,\quad t\leqq n.$$

Here the z_n are dimensionless constants which cannot easily be computed. It is therefore not sure, whether the series (3) converges even for finite a, but it is highly probable that it converges if $\delta=e^2/(hc)\cdot\lg\,(h/mca)<1$. One then would get $W=mc^2O(\delta)$ where $O(\delta)=1$ for a value of $\delta<1$. We then can define an electron radius in the same way as the classical radius e^2/mc^2 is defined, by putting the self-energy equal to mc^2. One obtains then roughly a value $a\sim h/(mc)\cdot\exp\,(-hc/e^2)$ which is about 10^{-58} times smaller than the classical electron radius. The "critical length" of the positron theory is thus infinitely smaller than usually assumed.

The situation is, however, entirely different for a particle with Bose statistics. Even the Coulombian part of the self-energy diverges to a first approximation as $W_{\text{el}}\sim e^2h/(mca^2)$ and requires a much larger critical length that is $a=(hc/e^2)^{\frac{1}{2}}\cdot h/(mc)$, to keep it of the order of magnitude of mc^2. This may indicate that a theory of particles obeying Bose statistics must involve new features at this critical length, or at energies corresponding to this length; whereas a theory of particles obeying the exclusion principle is probably consistent down to much smaller lengths or up to much higher energies.

II. The Charge Distribution of the Electron

The charge distribution in the neighborhood of an electron can be determined from the expression

$$G(\xi)=\int\rho(\mathfrak{r}-\xi/2)\rho(\mathfrak{r}+\xi/2)d\mathfrak{r}\,; \tag{4}$$

here $\rho(\mathfrak{r})$ is the charge density at the point $\mathfrak{r}$. $G(\xi)$ is the probability of finding charge simultaneously at two points in a distance ξ. If applied to a situation in which one electron alone is present, direct information can be drawn from this expression concerning the charge distribution in the electron itself. The charge density is given by

$$\rho(\mathfrak{r})=e\{\psi^*(\mathfrak{r})\psi(\mathfrak{r})\}-\sigma, \tag{5}$$

where $\psi(\mathfrak{r})$, the wave function, is a spinor with four components ψ_μ, $\mu=1, 2, 3, 4$. We write

$$\{\psi^*\psi\}=\sum_{\mu=1}^{4}\psi_\mu^*\psi_\mu$$

for the scalar product of two spinors. σ is the charge density of the unperturbed electrons in the negative energy states which is to be subtracted in the positron theory. In the one-electron theory σ is zero. The wave function ψ can be expanded in wave functions φ_q of the

[7] W. Pauli and V. Weisskopf, Helv. Phys. Acta **7**, 709 (1934).

stationary states q of a free electron:

$$\psi(\mathbf{r})=\sum_q a_q\varphi_q(\mathbf{r}). \tag{6}$$

The following relation holds for the φ_q

$$\{\varphi^*{}_q(\mathbf{r})\varphi_q(\mathbf{r})\}=1/V, \tag{7}$$

where V is the total volume of the system. We denote functions with positive energy values by φ_{+q} and with negative energy values by φ_{-q}.

We apply the method of quantized waves and consider the ψ's as operators acting on eigenfunctions $c(\cdots N_q\cdots)$ whose variables are the numbers N_q of electrons in different states q. If the ψ's are written in the form (6), the a's are operators which fulfill the well-known relations:

$$a_q{}^*a_q=N_q,\quad a_qa_q{}^*=1-N_q. \tag{8}$$

We now insert (6) into (5) and (5) into (4) and keep only terms which contain the products of two a's of the form (8) or the following combinations of four a's:

$$\begin{aligned} &a_q{}^*a_qa_{q'}{}^*a_{q'}=N_qN_{q'},\\ &a_q{}^*a_{q'}a_{q'}{}^*a_q=N_q(1-N_{q'}). \end{aligned} \tag{9}$$

All other combinations do not contribute to the expectation value $\bar{G}(\xi)$ of $G(\xi)$ because they have no diagonal elements. We obtain then

$$\bar{G}(\xi)=e^2\sum_q\sum_{q'}N_qN_{q'}+e^2\sum_q\sum_{q'}N_q(1-N_{q'})$$

$$\times\int\{\varphi_q{}^*(\mathbf{r}_1)\varphi_{q'}(\mathbf{r}_1)\}\{\varphi_{q'}{}^*(\mathbf{r}_2)\varphi_q(\mathbf{r}_2)\}d\mathbf{r}$$

$$-2\sigma e\sum_q N_q+\sigma^2V. \tag{10}$$

Here and in the following formulas we put $\mathbf{r}_1=\mathbf{r}-\boldsymbol{\xi}/2$, $\mathbf{r}_2=\mathbf{r}+\boldsymbol{\xi}/2$.

We first apply this expression to a single electron. We then put $\sigma=0$, and $N_{q_0}=1$ for $q=q_0$, $N_q=0$ for $q\neq q_0$:

$$\bar{G}(\xi)=e\sum_q\int\{\varphi_{q_0}{}^*(\mathbf{r}_1)\varphi_q(\mathbf{r}_1)\}\{\varphi_q{}^*(\mathbf{r}_2)\varphi_{q_0}(\mathbf{r}_2)\}d\mathbf{r}.$$

This expression can be evaluated by inserting the wave functions of a free electron. If q_0 is the state of an electron at rest, one obtains after replacing the sum over q by an integral over the momenta $\mathbf{p}$ of the states, the following result:

$$\bar{G}(\xi)=e^2\int d\mathbf{p}\frac{\exp i(\boldsymbol{\xi}\cdot\mathbf{p})/h}{8\pi^3h^3}=e^2\delta(\xi). \tag{11}$$

Thus in the one-electron theory, $\bar{G}(\xi)$ is equal to the δ-function.

We now apply (10) to the vacuum of the positron theory, that is, we set

$$N_{+q}=0,\quad N_{-q}=1,\quad \sigma=\sum_{-q}N_q.$$

It is easily seen, that the first, the third and the fourth term cancel each other. The terms remaining give

$$\bar{G}_{\text{vac}}(\xi)=e^2\sum_{+q}\sum_{-q'}\int\{\varphi_{-q'}{}^*(r_1)\varphi_{+q}(r_1)\}$$

$$\times\{\varphi_{+q}{}^*(r_2)\varphi_{-q'}(r_2)\}d\mathbf{r}.$$

The fact that this expression is different from zero and even infinite in the vacuum is closely connected with the charge fluctuations of the empty space which have been investigated by Heisenberg and Oppenheimer.[8] Heisenberg has shown that the charge fluctuations are infinite if the region in which they are measured is sharply limited. This result is due to the electron pairs produced when the charge is measured in a sharply defined region.

We are at present interested in the expression $\tilde{G}(\xi)$ corresponding to the charge distribution of one electron. This can be obtained by calculating $G_{\text{vac}+1}(\xi)$ for the state in which one electron in the state $+q_0$ is present ($N_{q_0}=1$ all other $N_{+q}=0$, $N_{-q}=1$), and by subtracting the effect of the vacuum $G_{\text{vac}}(\xi)$:

$$\tilde{G}(\xi)=G_{\text{vac}+1}(\xi)-G_{\text{vac}}(\xi)=$$

$$e^2\Big(\sum_{+q}-\sum_{-q}\Big)\int\{\varphi_{q_0}{}^*(r_1)\varphi_q(r_1)\}\{\varphi_q{}^*(r_2)\varphi_{q_0}(r_2)\}d\mathbf{r}.$$

If one inserts the actual solutions φ_q of Dirac's wave equation of the free electron, this expression can be readily evaluated. One obtains after replacing the sum by an integral as before:

$$\tilde{G}(\xi)=e^2mc^2\int\mathbf{dp}\frac{\exp i(\boldsymbol{\xi}\cdot\mathbf{p})/h}{8\pi^3h^3E(p)}; \tag{12}$$

[8] W. Heisenberg, Verh. d. Sächs. Akad. **86**, 317 (1934); J. R. Oppenheimer, Phys. Rev. **47**, 144 (1934).

here $E(p)=c(p^2+m^2c^2)^{\frac{1}{2}}$. This integral can be evaluated and gives:

$$\tilde{G}(\xi)=e^2\frac{mc}{h}\frac{1}{\xi}\frac{\partial}{\partial\xi}\frac{i}{2\pi}H_0^{(1)}\left(\frac{imc}{h}\cdot\xi\right).$$

$H_0^{(1)}(x)$ is the Hankel function of first kind; this function has a logarithmic singularity for $x=0$ and falls off exponentially for $x\gg1$. We obtain thus

$$\tilde{G}(\xi)=\begin{cases}\dfrac{e^2}{4\pi^2}\dfrac{mc}{h}\dfrac{1}{\xi^2} & \text{for } \xi\ll h/mc\\[2ex] e^2\cdot\left(\dfrac{mc}{h}\right)^2(h/2\pi^3mc\xi^3)^{\frac{1}{2}}\cdot e^{-mc\xi/h} & \\[2ex] & \text{for } \xi\gg h/mc.\end{cases}$$

This expression replaces the delta-function of the one-electron theory and indicates a spread of charge over a finite region of the order of h/mc. It is of interest to construct a charge density $\tilde{\rho}(r)$ for which

$$\int\tilde{\rho}(r_1)\tilde{\rho}(r_2)d\mathbf{r}=G(\xi).$$

This density is given by

$$\tilde{\rho}=e\int d\mathbf{p}\left(\frac{mc^2}{E(p)}\right)^{\frac{1}{2}}\frac{\exp i(\boldsymbol{\xi}\cdot\mathbf{p})/h}{8\pi^3h^3},$$

and for

$$r\ll h/mc:\ \tilde{\rho}\sim e\frac{mc}{h}2^{-5/2}\pi^{-3/2}r^{-5/2};$$

for $r\gg h/mc$, $\tilde{\rho}$ falls off exponentially.

In order to show that this "spread of charge" does not reduce the effect of a periodical field with a short wave-length, let us consider the operator

$$\int\rho(\mathbf{r})\exp i\mathbf{k}\cdot\mathbf{r}\,d\mathbf{r},$$

which represents an interaction energy between the charge and a field of wave number k. By inserting (5), this operator can be written in the form

$$e\sum_q a_q{}^*a_{q'},\quad \mathbf{p}_{q'}=\mathbf{p}_q+\mathbf{k}$$

and gives rise to transitions from any occupied state q' to any unoccupied state q. These transitions take place quite independently of the ratio of h/k to the linear dimensions h/mc of the spread of charge.

The energy W_{st} of the electrostatic field can be calculated directly from $\tilde{G}(\xi)$:

$$W_{st}=\tfrac{1}{2}\int\frac{\tilde{G}(\xi)}{|\xi|}d\xi.$$

The quadratic singularity of $\tilde{G}(\xi)$ at $\xi=0$ gives rise to a logarithmic divergence of W_{st}. By substituting (12) and by performing the integration over ξ first, we obtain the result

$$W_{st}=\frac{e^2}{4\pi^2}\int d\mathbf{p}\frac{mc^2}{h^3E_{(p)}p^2}$$

$$=\lim_{(P=\infty)}\frac{e^2}{\pi hc}mc^2\lg\frac{P+(P^2+m^2c^2)^{\frac{1}{2}}}{mc}\qquad(13)$$

or by putting $P=h/a$, where a is a length giving the "dimensions" of the electron, we get

$$W_{st}\sim\lim_{(a=0)}\frac{e^2}{\pi hc}m^2c^2\lg\frac{h}{mca}.$$

III. The Electromagnetic Field of the Electron

We calculate in this section the solenoidal part, E_s and H, of the electromagnetic field produced by the electron. It is given by

$$\mathbf{E}_s=-\frac{1}{c}\frac{\partial\mathbf{A}_s'}{\partial t},\quad \mathbf{H}=\operatorname{curl}\mathbf{A}'.$$

Here $\mathbf{A}_s'$ is the solenoidal part of the vector potential $\mathbf{A}$ which is given by

$$\mathbf{A}'(\mathbf{r},t)=\frac{1}{c}\int\frac{\mathbf{i}(\mathbf{r}',t-|\mathbf{r}-\mathbf{r}'|/c)}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}'.$$

A is primed to indicate that this field is produced by the electron. The current density $\mathbf{i}$ is defined by

$$i_x=ec\{\psi^*\alpha_x\psi\},\quad\text{etc.}$$

Here α_i are the well-known Dirac matrices. We consider in our approximation the wave func-

tions ψ to be the solutions of the wave equations of the field-free electron. If we expand the wave functions according to (6) we obtain

$$A_x'(\mathbf{r},t)=e\int d\mathbf{r}'\frac{\sum_q\sum_{q'}a_q^*a_{q'}\{\varphi_q^*(r't')\alpha_x\varphi_{q'}(r't')\}}{|\mathbf{r}-\mathbf{r}'|},$$

$$t'=t-\frac{|r-r'|}{c}.$$

The wave functions φ_q of the free electron are

$$\varphi_q=\frac{1}{V^{\frac{1}{2}}}u_q\exp\frac{1}{h}(i\mathbf{p}_q\cdot\mathbf{r}-iE_qt). \qquad (14)$$

p_q and E_q are momentum and energy of the state q, V is the volume of the space considered, u_q is a normalized spinor. We obtain then for A_x'

$$A_x(r,t)=\frac{2\pi eh^2c^2}{V}\sum_q\sum_{q'}\frac{\{u_q^*\alpha_xu_{q'}\}a_q^*a_{q'}}{E_qE_{q'}-m^2c^4-c^2\mathbf{p}_q\cdot\mathbf{p}_{q'}}$$
$$\times\exp i[(\mathbf{p}_{q'}-\mathbf{p}_q)\cdot\mathbf{r}-(E_{q'}-E_q)t]/h. \qquad (15)$$

The field strengths can immediately be computed from this expression. The time average of the magnetic field energy, U_{mag} is found to be

$$U_{\text{mag}}=\frac{1}{8\pi}\int(\text{curl }\mathbf{A}')^2d\mathbf{r}$$

$$=\frac{\pi}{2}e^2h^2c^4\sum_q\sum_{q'}(\mathbf{p}_q-\mathbf{p}_{q'})^2\cdot N_q(1-N_{q'})$$
$$\times\frac{\{u_q^*\alpha_su_{q'}\}\{u_{q'}^*\alpha_su_q\}}{[E_qE_{q'}-m^2c^4-c^2\mathbf{p}_q\cdot\mathbf{p}_{q'}]^2} \qquad (16)$$

if one uses the relations (9). α_s is the component of $\boldsymbol{\alpha}$ which is perpendicular to $\mathbf{p}_q-\mathbf{p}_{q'}$. We apply this expression first to a single electron at rest ($N_q=1$, $q=q_0$; $N_q=0$, $q\neq q_0$):

$$U_{\text{mag}}=\frac{\pi}{2}e^2h^2\sum_q p_q{}^2\frac{\{u_{q_0}^*\alpha_su_q\}\{u_q^*\alpha_su_{q_0}\}}{m^2(E_q-m^2c^4)^2}.$$

One obtains after averaging over the spin directions and replacing the sum by an integral:

$$U_{\text{mag}}=\pi\frac{e^2h^2}{m^2c^2}\int\frac{d\mathbf{p}}{8\pi^3h^3}=\frac{1}{2\pi}\lim_{(a=0)}\frac{e^2h^2}{3m^2c^2a^3}. \qquad (17)$$

This divergent expression corresponds to the field energy of a magnetic dipole density concentrated in a sphere of infinitely small radius a. If one now calculated expression (16) for the vacuum of the positron theory ($N_{+q}=0$, $N_{-q}=1$) one obtains a highly divergent expression which represents the magnetic field energy $U_{\text{mag}}(\text{Vac.})$ produced by the current fluctuations of the vacuum. We are interested in the field energy of one electron at rest which is obtained by calculating $U_{\text{mag}}(\text{Vac.}+1)$ for the state ($N_{q_0}=1$, $p_{q_0}=0$, $E_{q_0}=mc^2$ all other $N_{+q}=0$, $N_{-q}=1$) and by subtracting the effect of the vacuum:

$$\tilde{U}_{\text{mag}}=U_{\text{mag}}(\text{Vac.}+1)-U_{\text{mag}}(\text{Vac.})$$
$$=2\pi e^2h^2\Big(\sum_{+q}-\sum_{-q}\Big)p_q{}^2\frac{\{u_{q_0}^*\alpha_su_q\}\{u_q^*\alpha_su_{q_0}\}}{m^2(E_q-m^2c^4)^2}.$$

Averaging over the spin directions gives:

$$U_{\text{mag}}=\pi\frac{e^2h^2}{m}\int\frac{d\mathbf{p}}{8\pi^3h^3E(p)}$$

$$=\frac{1}{2\pi}\frac{e^2}{mch}\lim_{(P=\infty)}\left[PP_0-\frac{m^2c^2}{2}\lg\frac{P+P_0}{mc}\right]. \qquad (18)$$

Here P_0 is defined by $P_0=(P^2+m^2c^2)^{\frac{1}{2}}$. This quadratically divergent expression is just what one would expect for the field energy of the magnetic dipole density of the electron if this density is equal to the charge distribution calculated in the previous section.

In order to show directly, that this magnetic field is equal to the "static" field of the magnetic moment of the spin, we consider the magnetic polarization (dipole density) $\mathbf{M}$,

$$M_z=\frac{eh}{2mc}\{\psi^*\beta\alpha_y\alpha_z\psi\},$$

and calculate the function

$$J(\xi)=\int\mathbf{M}(r-\xi/2)\cdot\mathbf{M}(r+\xi/2)d\mathbf{r},$$

which corresponds to $G(\xi)$ (see (4)) and which provides information about the "spin distribution" in the electron. If one evaluates this integral by the method which is used in Section II,

one obtains

$$J(\xi)=\frac{3}{4}\frac{h^2}{m^2c^2}G(\xi),$$

which shows[9] that the spin is distributed in exactly the same way as the charge. The magnetic field energy of this distribution is given by

$$U_{\text{mag}}=\tfrac{1}{2}\int\frac{\text{div. }\mathbf{M}(r_1)\ \text{div. }\mathbf{M}(r_2)}{|\xi|}d\mathbf{r}$$

and can be evaluated by the methods used for the other calculations in this section and leads to the expressions (17) for the one-electron theory and to (18) for the positron theory. Hence we are allowed to consider them as the field energy of the magnetic moment.

The energy of the solenoidal electric field strength is given by

$$U_{\text{el}}=\frac{1}{8\pi c^2}\int\left(\frac{\partial A_s}{\partial t}\right)^2 d\mathbf{r}$$

$$=\frac{\pi}{2}\frac{e^2h^2}{c^2}\sum_q\sum_{q'}(E_{q'}-E_q)^2\frac{\{u_q{}^*\alpha_s u_{q'}\}\{u_{p'}{}^*\alpha_s u_q\}}{[E_qE_{q'}-m^2c^4-c^2\mathbf{p}_q\cdot\mathbf{p}_{q'}]^2}N_q(1-N_{q'}).$$

Applied to a single electron at rest, this expression gives

$$U_{\text{el}}=\pi\frac{e^2h^2}{m^2c^2}\int\frac{\mathbf{dp}}{8\pi^3h^3}=U_{\text{mag}}. \qquad (19)$$

Applied to an electron at rest in positron theory, one obtains, however,

$$\tilde{U}_{\text{el}}=U_{\text{el}}(\text{Vac.}+1)-U_{\text{el}}(\text{Vac.})$$

$$=\frac{2\pi e^2h^2}{c^2}\left(\sum_{+q}-\sum_{-q}\right)(E_q-mc^2)^2\times\frac{\{u_{q_0}{}^*\alpha_s u_q\}\{u_q{}^*\alpha_s u_{q_0}\}}{m^2(E_q-m^2c^4)^2}$$

and then

$$\tilde{U}_{\text{el}}=-\pi\frac{e^2h^2}{m}\int\frac{\mathbf{dp}}{8\pi^3h^3E(p)}=-\tilde{U}_{\text{mag}}. \qquad (20)$$

The interpretation of this result is given in Section I.

[9] The factor $\frac{3}{4}$ is $s(s+1)$ for $s=\frac{1}{2}$.

IV. The Self-Energy of the Electron

The self-energy is calculated in this section by means of a method which is different from the usual perturbation method, in order to outline the physical significance of the different terms. It is similar to the method applied to this problem previously by the author.[5]

The Hamiltonian of a system of charged particles and their electromagnetic field can be written in the form

$$\mathfrak{H}=\tfrac{1}{2}\int\left[\frac{1}{c^2}\left(\frac{\partial\mathbf{A}}{\partial t}\right)^2+\sum_{i,\,k=1}^{3}\left(\frac{\partial A_i}{\partial x_k}\right)^2-\frac{1}{c^2}\left(\frac{\partial\phi}{\partial t}\right)^2-(\text{grad. }\phi)^2\right]d\mathbf{r}$$

$$+\int\left(\rho\phi-\frac{1}{c}\mathbf{i}\cdot\mathbf{A}\right)d\mathbf{r}$$

$$+c\sum_i\int\{\psi_i{}^*(\boldsymbol{\alpha}\cdot\mathbf{p}_i+\beta mc)\psi_i\}d\mathbf{r}.$$

The summation in the last term is performed over all particles. The solutions of this Hamiltonian are restricted by the condition

$$\partial\phi/\partial t+c\ \text{div. }\mathbf{A}=0.$$

By introducing the field strengths instead of the potentials we obtain the expression

$$\mathfrak{H}=\frac{1}{8\pi}\int(E^2+H^2)d\mathbf{r}-\frac{1}{4\pi}\int\phi\cdot\text{div. }E d\mathbf{r}$$

$$+\int\left(\rho\phi-\frac{1}{c}\mathbf{i}\cdot\mathbf{A}\right)d\mathbf{r}$$

$$+c\sum_i\int\{\psi_i{}^*(\boldsymbol{\alpha}\cdot\mathbf{p}_i+\beta mc)\psi_i\}d\mathbf{r}.$$

This is equal to (1) if one uses the relations div. $E=4\pi\rho$ and

$$T=c\sum_i\int\{\psi_i{}^*[\boldsymbol{\alpha}\cdot(\mathbf{p}-(e/c)\mathbf{A})+\beta mc]\psi\}d\mathbf{r}.$$

The interaction energy

$$\int(\rho\phi-(1/c)\mathbf{i}\cdot\mathbf{A})=eH'$$

between matter and field contains the electronic charge e explicitly as a linear factor, so that in the above notation H' is explicitly independent of e; $\partial H'/\partial e=0$. Let us consider the energy W_s of a stationary state s of this Hamiltonian. If the electronic charge e is increased by de the Hamiltonian gets the additional term $de\,H'$. According to perturbation theory the increase of W_s is $de\langle H'\rangle_{Av}$, where $\langle H'\rangle_{Av}$ is the time average of H' in the state s, assuming that the electronic charge has its original value e. We therefore get

$$W_s=W_s^{(0)}+\int_0^e\langle H'(e)\rangle_{Av}de, \tag{21}$$

where $W_s^{(0)}$ is the value of the energy for $e=0$. We now expand $H'(e)$ in a power series of e:

$$H'(e)=H'^{(0)}+eH'^{(1)}+\cdots$$

and get from (21)

$$W_s=W_s^{(0)}+e\langle H'^{(0)}\rangle_{Av}+\frac{e^2}{2}\langle H'^{(1)}\rangle_{Av}+\cdots.$$

The second term is zero since W_s cannot depend upon the sign of e, and we obtain by neglecting all terms containing e in a higher power than the second

$$W'=W_s-W_s^{(0)}$$

$$=\frac{e^2}{2}\langle H'^{(1)}\rangle_{Av}=\frac{e}{2}\langle H'\rangle_{Av}=\tfrac{1}{2}\int(\langle\rho\phi\rangle_{Av}-\frac{1}{c}\langle\mathbf{i}\cdot\mathbf{A}\rangle_{Av})d\mathbf{r}.$$

The potentials can be split into two parts

$$\mathbf{A}=\mathbf{A}_0+\mathbf{A}',\qquad \phi=\phi_0+\phi'.$$

$\mathbf{A}_0$ and ϕ_0 are the potentials of the field when no electron is present. In empty space, A_0 is the potential of the zero-point oscillations and $\phi_0=0$. A' and ϕ' is the field produced by the electron. We then get

$$W'=\tfrac{1}{2}\int\left(\langle\rho\phi'\rangle_{Av}-\frac{1}{c}\langle\mathbf{i}\cdot\mathbf{A}'\rangle_{Av}\right)d\mathbf{r}$$

$$-\frac{1}{2c}\int\langle\mathbf{i}\cdot\mathbf{A}_0\rangle_{Av}d\mathbf{r}.$$

The first term can be transformed by means of the relations

$$\Delta\phi'-\frac{1}{c^2}\ddot{\phi}'=-4\pi\rho,\quad \Delta\mathbf{A}'-\frac{1}{c^2}\ddot{\mathbf{A}}'=-4\pi\mathbf{i},$$

and we obtain

$$W'=\frac{1}{8\pi}\int(\langle E'^2\rangle_{Av}-\langle H'^2\rangle_{Av})d\mathbf{r}-\frac{1}{2c}\int\langle\mathbf{i}\cdot\mathbf{A}_0\rangle_{Av}d\mathbf{r}.$$

We consider now the state s to be the state of an electron at rest. The charge dependent part W' of the self-energy can then be written in the form

$$W'=W_{st}+W_{sp}+W_{fluct}. \tag{22}$$

Here

$$W_{st}=\frac{1}{8\pi}\int E_{st}^2d\mathbf{r}$$

is the static field energy of the irrotational field E_{st}; W_{sp} is defined by

$$W_{sp}=\frac{1}{8\pi}\int(\langle E_s'^2\rangle_{Av}-\langle H'^2\rangle_{Av})d\mathbf{r}=U_{el}-U_{mag}. \tag{23}$$

It contains the contribution of the field produced by the spin and is calculated in the previous section; W_{fluct} is the energy produced by the fluctuations of the radiation field,

$$W_{fluct}=-\frac{1}{2c}\int\mathbf{i}\cdot\mathbf{A}_0d\mathbf{r}.$$

In order to calculate W_{fluct}, we divide $\mathbf{i}$ into two parts $\mathbf{i}=\mathbf{i}_0+\mathbf{i}'$, where $\mathbf{i}_0$ is the current density for the field-free case. The term $\int\mathbf{i}_0\cdot\mathbf{A}_0d\mathbf{r}$ vanishes when averaged over the time because of the absence of phase relations between $\mathbf{i}_0$ and $\mathbf{A}_0$. The remaining term

$$W_{fluct}=-\frac{1}{2c}\int\mathbf{i}'\cdot\mathbf{A}_0d\mathbf{r}$$

can be evaluated as follows:

$\mathbf{i}'$ is to a first approximation given by

$$\mathbf{i}'=\{\psi_0^*\boldsymbol{\alpha}\psi_1\}+\{\psi_1^*\boldsymbol{\alpha}\psi_0\}. \tag{24}$$

Here $\psi_0=\Sigma a_q\varphi_q$ is the wave function unperturbed by the field and ψ_1 is the first approxima-

tion of the perturbed wave function:

$$\psi_1 = \sum a_q \varphi_q'.$$

φ_q' can be calculated by means of the ordinary perturbation method. The interaction energy with an arbitrary field $\mathbf{A}$ is given by

$$-e\alpha\mathbf{A} = -e\alpha\sum_k [A_k^{(+)} \exp i(\mathbf{k}\cdot\mathbf{r}+c|k|t) + A_k^{(-)} \exp -i(\mathbf{k}\cdot\mathbf{r}+c|k|t)].$$

The sum is taken over all wave numbers $\mathbf{k}$ of an orthogonal system of plane waves in the volume V. If we write the wave functions φ_q of the free electron in the form (14), we get for φ_q'

$$\varphi_q' = e\sum_{q'} \varphi_{q'} \left[\frac{\{u_{q'}^* \alpha u_q\} A_k^{(+)}}{E_q - E_{q'} + c|k|} e^{+ic|k|t} + \frac{\{u_{q'}^* \alpha u_q\} A_k^{(-)}}{E_q - E_{q'} - c|k|} e^{-ic|k|t} \right], \quad h\mathbf{k} = \mathbf{p}_{q'} - \mathbf{p}_q.$$

This expression is introduced into (24) and gives for the current

$$i_z' = \sum_q \sum_{q'} N_q \{u_q^* \alpha_z u_{q'}\} \{u_{q'}^* \boldsymbol{\alpha} u_q\} \times \left[\frac{\mathbf{A}^{(+)} \exp i(\mathbf{k}\cdot\mathbf{r}+c|k|t)}{E_q - E_{q'} + c|k|} + \frac{\mathbf{A}^{(-)} \exp -i(\mathbf{k}\cdot\mathbf{r}+c|k|t)}{E_q - E_{q'} - c(k)} \right] + \text{conj.}$$

We have retained only terms containing products (8) or (9). It is seen from this expression, that i' is the same in the one-electron theory and in the positron theory. In the latter case we have to consider $\bar{i} = i'(\text{Vac.}+1) - i'(\text{Vac.})$. The actual value can easily be evaluated for an electron at rest. One obtains

$$\mathbf{i}' = -\frac{1}{V}\frac{e^2}{mc}\mathbf{A},$$

which is immediately understood as the forced vibrations of a point charge under the action of an oscillating field $\mathbf{A}$. W_{fluct} is directly obtained if one replaces $\mathbf{A}$ by the field fluctuations $\mathbf{A}_0$ of the vacuum:

$$W_{\text{fluct}} = -\frac{1}{2c}\int \mathbf{i}' \cdot \mathbf{A}_0 \mathbf{dr} = \frac{e^2}{2mc^2} \langle A_0^2 \rangle_{Av}.$$

We have

$$\langle A_0^2 \rangle_{Av} = \frac{1}{\pi^2}\int \frac{ch}{|k|} \mathbf{dk} \tag{25}$$

and we get finally, on replacing $|k|$ by p/h,

$$W_{\text{fluct}} = \frac{e^2}{2\pi^2 mc^2}\int \frac{c\mathbf{dp}}{h|p|} = \frac{e^2}{\pi hc}\frac{1}{m}\lim_{(P=\infty)} P^2.$$

We now collect the results obtained for the other parts of the self-energy of an electron at rest. The one-electron theory gives [Eqs. (11), (23), (19)]

$$W_{\text{st}} = \lim_{(a=0)} \frac{e^2}{a}, \quad W_{\text{sp}} = U_{\text{el}} - U_{\text{mag}} = 0.$$

The positron theory gives (Eqs. (13), (23), (20))

$$W_{\text{st}} = \lim_{(P=\infty)} \frac{e^2}{\pi hc} mc^2 \lg \frac{P+P_0}{mc},$$

$$W_{\text{sp}} = U_{\text{el}} - U_{\text{mag}} = -\frac{e^2}{\pi hc}\frac{1}{m}\lim_{(P=\infty)} \left[PP_0 - \frac{m^2c^2}{2} \lg \frac{P+P_0}{mc} \right].$$

W_{sp} is partly balanced by W_{fluct}. The total self-energy in positron theory is then given by

$$W' = \frac{3}{2\pi}\frac{e^2}{hc} mc^2 \lim_{(P=\infty)} \lg \frac{P+P_0}{mc} + \text{finite terms.} \tag{26}$$

The self-energy of a free electron in motion can be obtained by a Lorentz transformation from (26). The direct calculation from the above methods is ambiguous because it leads to a difference of terms, each of which diverges quadratically. The factor of the logarithmically divergent difference of these terms depends essentially on the way in which the infinite terms are subtracted. The calculation of the self-energy of an electron at rest is not so much exposed to these ambiguities because of the spherical symmetry of the problem, which suggests only one natural way of subtracting two divergent integrals over the momentum space, namely, the subtraction of the contributions of concentric spherical shells around the center. It must be expected, that the value of the self-energy of a moving electron can only be covariant to the value (26) if one performs the subtraction appropriately. This is why the expressions for the self-energy obtained in reference 5 are apparently not relativistically covariant.

V. The Self-Energy of a Particle Obeying the Bose-Statistics

It has been shown that the quantization of the scalar wave equation of Klein and Gordon leads to a theory of elementary particles with Bose statistics and charges of both signs. The theory includes a description of pair creation and of all related phenomena. The quantitative results are not very different from the results of Dirac's positron theory. The formalism has been recently applied to particles with intrinsic angular momentum. It will be shown here that the calculation of the self-energy, however, gives results quite different from the positron theory. The energy of the electrostatic field of the electron is found to be more strongly divergent than in the classical theory; the energy of the radiation field diverges quadratically and is equal to the corresponding energy of a single electron in the one-electron theory. The qualitative arguments for this behavior are given in I. The following calculation is based on the formulas derived elsewhere.[7]

The operator of the charge density is given by

$$\rho = ie(\psi^*\pi^* - \psi\pi),$$

where ψ is the wave function and π its conjugate operator:

$$\pi = h\frac{\partial\psi^*}{\partial t},$$

$$\psi(r)\pi(r') - \pi(r')\psi(r) = \delta(r-r').$$

We introduce new variables by means of

$$\psi = \frac{1}{V^{\frac{1}{2}}}\sum_k q(\mathbf{k})\exp i\mathbf{k}\cdot\mathbf{r},$$
$$\pi = \frac{1}{V^{\frac{1}{2}}}\sum_k p(\mathbf{k})\exp -i\mathbf{k}\cdot\mathbf{r}. \tag{27}$$

Here $1/V^{\frac{1}{2}}\cdot\exp i\mathbf{kr}$ form a set of orthogonal functions in the volume V.

We further introduce

$$p(k) = \left(\frac{E(k)}{2}\right)^{\frac{1}{2}}(a^*(\mathbf{k}) + b(\mathbf{k})),$$
$$q(k) = -i\left(\frac{1}{2E(k)}\right)^{\frac{1}{2}}(a^*(\mathbf{k}) - b(\mathbf{k})), \tag{28}$$
$$E(k) = c(h^2k^2 + m^2c^2)^{\frac{1}{2}},$$

and obtain for

$$\rho(\mathbf{s}) = \frac{1}{V}\int \rho(\mathbf{r})\exp(-i\mathbf{s}\cdot\mathbf{r})\mathbf{dr}$$

the following expression:

$$\rho(\mathbf{s}) = \frac{e}{2V}\sum_k \frac{E(k)+E(l)}{[E(l)E(k)]^{\frac{1}{2}}}[a^*(\mathbf{k})a(\mathbf{l}) - b^*(\mathbf{k})b(\mathbf{l})]$$
$$+\frac{E(k)-E(l)}{[E(k)E(l)]^{\frac{1}{2}}}[a(\mathbf{k})b(\mathbf{l}) - a^*(\mathbf{k})b^*(\mathbf{l})], \tag{29}$$

where $\mathbf{l} = \mathbf{k} + \mathbf{s}$.

The $a(\mathbf{k})$ and $b(\mathbf{k})$ fulfill the relations

$$a^*(\mathbf{k})a(\mathbf{k}) = N(\mathbf{k}), \qquad b^*(\mathbf{k})b(\mathbf{k}) = M(\mathbf{k}),$$
$$a(\mathbf{k})a^*(\mathbf{k}) = 1 + N(\mathbf{k}), \qquad b(\mathbf{k})b^*(\mathbf{k}) = 1 + M(\mathbf{k}).$$

The $N(\mathbf{k})$'s are the numbers of positrons, the $M(\mathbf{k})$'s the number of negatrons in the states with the wave vector $\mathbf{k}$. The electrostatic self-energy is given by:

$$W_{st} = \tfrac{1}{2}\int \frac{\rho(\mathbf{r}-\xi/2)\rho(r+\xi/2)}{|\xi|}\mathbf{drd}\xi$$
$$= 2\pi V\sum_s \frac{\rho(-\mathbf{s})\rho(\mathbf{s})}{s^2}.$$

We obtain by introducing (29) and retaining only the diagonal terms:[10]

$$W_{st} = \frac{\pi e^2}{2V}\sum_k\sum_s \frac{1}{s^2}\Big[\frac{(E(k)+E(l))^2}{E(k)E(l)}$$
$$\times(N(\mathbf{k})[N(\mathbf{l})+1] + M(\mathbf{k})[M(\mathbf{l})+1])$$
$$+\frac{(E(\mathbf{k})-E(\mathbf{l}))^2}{E(k)E(l)}$$
$$\times(N(\mathbf{k})M(\mathbf{l}) + [N(\mathbf{k})+1][M(\mathbf{l})+1])\Big].$$

This expression does not vanish for the vacuum ($N(\mathbf{k}) = M(\mathbf{k}) = 0$ for every $\mathbf{k}$). We calculate the difference

$$\bar{W}_{st} = W_{st}(\text{Vac.}+1) - W_{st}(\text{Vac.})$$

[10] The term $s=0$ is omitted. It can easily be shown that this term does not contribute for $V\to\infty$.

and assume that the particle is at rest:

$$\bar{W}_{st}=\frac{e^2h}{16mc}\int ds\frac{1+2m^2c^2/h^2s^2}{(s^2+m^2c^2/h^2)^{\frac{1}{2}}}$$
$$=e^2\frac{h}{mc}\frac{\pi}{4}\lim_{(P=\infty)}\left[\frac{P^2}{h^2}+\frac{m^2c^2}{h^2}\lg\frac{P+P_0}{mc}\right].$$

This is an expression which diverges quadratically. By putting $P=h/a$ one obtains

$$\bar{W}_{st}\leqq\frac{\pi}{4}mc^2\left(\frac{h}{mca}\right)^2.$$

We now show that the particle does not produce a solenoidal electric or magnetic field in the approximation considered here. The current density is given by Eq. (43) of reference 7:

$$\mathbf{i}=\mathbf{i}_1+\mathbf{i}_2,$$
$$\mathbf{i}_1=ihce(\psi\ \text{grad.}\ \psi^*-\psi^*\ \text{grad.}\ \psi),$$
$$\mathbf{i}_2=-2e^2\mathbf{A}\psi^*\psi.$$

We introduce the new variables (27) into i_1 and get

$$\mathbf{i}_1=\sum_{\mathbf{s}}i(\mathbf{s})\exp i\mathbf{r}\cdot\mathbf{s},$$

$$i(\mathbf{s})=\tfrac{1}{2}hce\sum_{\mathbf{k}}\frac{\mathbf{k}+1}{[E(k)E(l)]^{\frac{1}{2}}}(a^*(\mathbf{k})a(\mathbf{l})+b(\mathbf{k})b^*(\mathbf{l})$$
$$-a(\mathbf{k})b(\mathbf{l})-b(\mathbf{k})a(\mathbf{l})),\quad \mathbf{l}=\mathbf{k}+\mathbf{s}.$$

$i_1(\mathbf{s})$ is proportional to $\mathbf{k}+\mathbf{l}$ and $\mathbf{i}_1$ is therefore irrotational for all transitions which start or end with a particle at rest. ($\mathbf{k}+\mathbf{l}$ is parallel to $\mathbf{s}$ if $\mathbf{k}=0$ or $\mathbf{l}=0$). Thus $\mathbf{i}_1$ does not produce a solenoidal field. $\mathbf{i}_2$ is proportional to e^2 so that its field does not come into consideration. The particle does not give rise to a solenoidal field as long as it is at rest since it has no magnetic spin moment.

The remaining term of the self-energy is

$$W_{fluct}=-\frac{1}{2c}\int\mathbf{i}'\cdot\mathbf{A}_0\mathbf{dr}. \qquad (22)$$

$\mathbf{i}'$ is defined as the current density produced by the field $\mathbf{A}_0$. Here the first part $\mathbf{i}_1'$ can be shown to be again irrotational. The integral over the product of $\mathbf{i}_1'$ and the solenoidal vector $\mathbf{A}_0$ vanishes. The second part $\mathbf{i}_2'$ is in the required approximation directly given by

$$i_2'=-2e^2\mathbf{A}_0\psi^*\psi.$$

We obtain then

$$W_{fluct}=e^2\int\mathbf{A}_0^2\psi^*\psi\mathbf{dr}.$$

By introducing the new variables and retaining only diagonal elements we find

$$W_{fluct}=\frac{e^2}{2}\langle A_0^2\rangle_{Av}\sum_{\mathbf{k}}\frac{N(\mathbf{k})+M(\mathbf{k})+1}{E(k)}$$

and finally

$$\bar{W}_{fluct}=W_{fluct}(\text{Vac.}+1)-W_{fluct}(\text{Vac.})$$
$$=\frac{e^2}{2mc^2}\langle A_0^2\rangle_{Av}.$$

This is identical with the corresponding expression for the Dirac electron.

VI. The Higher Approximations of the Self-Energy in the Theory of the Positron

It will be proved in this section that the successive approximations of the self-energy of the electron vanish in the limit $m\to0$. Furthermore it is shown that the divergence of the self-energy is logarithmic in every approximation.[11]

We consider the total system containing the electrons and the radiation field and calculate the energy $W(s)$ of the state s of this system. $W(s)$ can be expanded in a series of approximations $W(s)=\Sigma_n W^{(n)}(s)$ corresponding to an expansion in powers of the parameter e^2/hc. Since this procedure does not give zero for the vacuum in Dirac's positron theory, the self-energy of the electron must be defined as the difference between the energy $W(\text{Vac.}+1)$ of the state in which one electron is present and the energy $W(\text{Vac.})$ of the vacuum alone. We confine ourselves to the calculation of the electrodynamic

[11] Recently A. Mercier (Helv. Phys. Acta 12, 55 (1938)) has treated the same problem and has obtained a higher divergence. As he does not compute the numerical factors of the divergent expressions, he cannot exclude a factor zero for the highest divergent terms. The following considerations, however, show that the highest nonvanishing terms diverge only logarithmically.

self-energy. The calculation of the electrostatic energy and the mixed terms in higher approximation can be made along the same lines.

We now consider the detailed form of the nth approximation $W_s^{(n)}$ of the energy of the state s. $W_s^{(n)}$ is a sum of terms containing a product of $2n$ matrix elements of the interaction energy which correspond to consecutive transitions of the total system from one state to another starting from the state s and returning to it. The terms have denominators that are products of energy differences between the original state s and intermediate states.

The self-energy of the electron $\tilde{W}^{(n)}$ is given by the difference

$$\tilde{W}^{(n)} = W^{(n)}(\text{Vac.}+1) - W^{(n)}(\text{Vac.}). \quad (30)$$

We now prove that $\tilde{W}^{(n)}=0$ for $m=0$ by comparing it with the self-energy of a positron $\tilde{W}'^{(n)}$ in the same state:

$$\tilde{W}'^{(n)} = W^{(n)}(\text{Vac.}-1) - W^{(n)}(\text{Vac.}) = \tilde{W}^{(n)}, \quad (31)$$

which is equal to the self-energy of the electron. There is no loss of generality if we confine our considerations to the self-energy of an electron at rest. The state (Vac.+1) is then specified by: every negative energy state and the lowest positive state s_{+0} occupied; (Vac.−1) means: every negative energy state except the highest one s_{-0} is occupied. We now show that

$$W^{(n)}(\text{Vac.}+1)+W^{(n)}(\text{Vac.}-1) = 2W^{(n)}(\text{Vac.}) \quad \text{for} \quad m=0. \quad (32)$$

Comparing $W^{(n)}(\text{Vac.}+1)$ and $W^{(n)}(\text{Vac.}-1)$ with $W^{(n)}(\text{Vac.})$ we notice:

(a) $W^{(n)}(\text{Vac.}-1)$ lacks all terms containing transitions of the electron in the state s_{-0}.

(b) $W^{(n)}(\text{Vac.}+1)$ contains additional terms from transitions of the additional electron in s_{+0}.

(c) $W^{(n)}(\text{Vac.}-1)$ contains additional terms from transitions of one of the vacuum electrons into the empty state s_{-0}.

(d) $W^{(n)}(\text{Vac.}+1)$ lacks terms containing transitions of the vacuum electrons into the state s_{+0} of the additional electron.

We now prove that the missing terms of (a) and (d) are in the limit $m\to 0$ identical with the additional terms of (b) and (c), respectively. The only difference between the pair (a), (b) and the pair (c), (d) consists in the fact that the specified transitions start from a positive or from a negative state, respectively. This fact does not affect the energy differences in the denominators in the limit $m\to 0$, as in this limit the energies of s_{+0} and s_{-0} are equal. It remains to show that the numerators are also unchanged. This can be seen in the following way: the transition matrix elements appearing in the denominator belong to a chain of consecutive transitions starting from and returning to the initial state of the system. Thus the transition elements from or to s_{+0} or s_{-0} belong to a chain of transitions of an electron starting from the state s_{+0} (or from s_{-0}, respectively) and returning to this state. The transition elements form the product:

$$P_\pm = \{\psi^*(\pm 0)H_1\psi(p_1)\}\{\psi^*(p_1)H_2\psi(p_2)\}\cdots \times\{\psi^*(p_{n-1})H_n\psi(\pm 0)\}. \quad (33)$$

Here H_i is the interaction energy with the light quantum which is emitted or absorbed with the ith transition and $\psi(p_i)$ is the wave function of the electron with the momentum p_i which performs the transition. $\psi(\pm 0)$ is the wave function of the state s_{+0} or s_{-0}, respectively. After averaging over the two spin states for every momentum p_i, this product can be written as trace of the following matrix:

$$P_\pm = \text{Trace}\left[H_1\cdot\tfrac{1}{2}\left(1+\frac{\boldsymbol{\alpha}\cdot\mathbf{p}_1+\beta mc}{E_1}\right)\cdot H_2 \cdot\tfrac{1}{2}\left(1+\frac{\boldsymbol{\alpha}\cdot\mathbf{p}_2+\beta mc}{E_2}\right)\cdots H_n\tfrac{1}{2}(1\pm\beta)\right].$$

Here $E_i=\pm(p_i^2+m^2c^2)^{\frac{1}{2}}$ is the energy of the wave function $\psi(p_i)$. If we now go to the limit $m\to 0$, all terms with β disappear except those containing the β in the last parenthesis. (H_i does not contain the operator β.)[12] Since the trace of every expression containing α's but only one β is zero, we are allowed to omit the last also, and it becomes evident that $P_+=P_-$. From (30), (31)

[12] This conclusion does not hold if one or more of the p_i's are $\sim mc$. As $\tilde{W}^{(n)}$ is an integral over all possible intermediate momenta p_i in the volume V, the terms with one or more $p_i\sim mc$ contain one or more factors $(mc)^3$ so that these contributions should be neglected.

and (32) it follows directly that

$$\tilde{W}^{(n)}=0 \quad \text{for} \quad m=0.$$

We now prove that this result infers a logarithmical divergence of $\tilde{W}^{(n)}$ for finite m. The terms of $\tilde{W}^{(n)}$ contain transitions in which n light quanta are emitted and absorbed. We write $W^{(n)}$ as an integral over the wave vectors $\mathbf{k}_1 \cdots \mathbf{k}_n$ of these light quanta:

$$\tilde{W}^{(n)}=\int_0^{P/h} \mathbf{dk}_1 \cdots \int_0^{P/h} \mathbf{dk}_n F(k_1 \cdots k_n) \quad (34)$$

and integrate to a finite limit $P/h \gg mc/h$ in order to make $\tilde{W}^{(n)}$ finite. The properties of $F(k_1 \cdots k_n)$ are very simple for all $p_i \gg mc$ where p_i are as above the momenta of the electrons which change their states in the transitions to intermediate states (except of course the momentum $p_0=0$ of the electron under consideration in its initial state). It can be shown that replacing every k_l $(l=1 \cdots n)$ by $k_l'=xk_l$ gives

$$F(k_1' \cdots k_n')=x^N F(k_1 \cdots k_n) \quad \text{if} \quad p_i \gg mc,$$

where N is an integer. This result may be understood in the following way. By means of the substitution $k_l'=xk_l$ the momenta p_i of the excited electrons in the intermediate states are also multiplied by a since the p_i are sums or differences of the momenta of the absorbed or emitted light quanta. If now $k_l \gg mc$, all momenta p_i involved are large compared to mc, and the corresponding energies can be replaced by $c|p_i|$. This neglect of mc compared to p_i has as consequence that all energy differences E_s-E_i between the initial state and the intermediate states are multiplied by x if one replaces k_l by k_l'. Since all terms of $\tilde{W}^{(n)}$ have $2n-1$ energy differences in the denominator, the latter is proportional to x^{2n-1}. The numerators consist of n matrix elements which form expressions like (33). From the fact that every H_i is proportional to k_i^{-1} and from the structure of (33) it follows that the numerators are sums of terms proportional to x^{-n}, $x^{-n-1} \cdots x^{-2n}$. We get, therefore,

$$F(k_1' \cdots k_n')=\sum_{z=3n-1}^{z=5n-1} c_z x^{-z} \quad \text{for} \quad p_i \gg mc.$$

Not all of these c_z need to be different from zero. If c_ξ is the first coefficient different from zero, we can write

$$F(k_1' \cdots k_n')=c_\xi x^{-\xi} \leqq x^{-\xi} F(k_1 \cdots k_n) \quad (35)$$

because we certainly can neglect the terms with $z>\xi$, which are smaller by the ratio $\sim(mc/p_i)^{z-\xi}$. We are now interested in the function $\tilde{W}^{(n)}(P)$. The relation (35) is not valid in the entire region of integration in (34). The regions in which the conditions $p_i \gg mc$ are not fulfilled are restricted to certain small areas in the $3n$-dimensional space of the wave vectors $\mathbf{k}_1 \cdots \mathbf{k}_n$. The contributions of these areas can be neglected for $P \gg mc$. We therefore get

$$\tilde{W}^{(n)}(xP) \leqq \int_0^{P/h} \mathbf{dk}_1' \cdots \int_0^{P/h} \mathbf{dk}_n' F(k_1' \cdots k_n')$$
$$\leqq x^{3n-\xi} \cdot W'^{(n)}(P) + \text{smaller terms.}$$

The additional smaller terms come from the regions of integration in which these considerations are not applicable and from the neglected terms. From this relation follows, that

$$\tilde{W}^{(n)} \leqq c\left(\frac{e^2}{hc}\right)^n \cdot \frac{P^N}{(mc)^N} \cdot \left(\lg \frac{P}{mc}\right)^t \cdot mc^2 + \text{smaller terms.}$$

Here $N=3n-\xi$, c is a numerical factor and $0 \leqq t \leqq n$ because any of the n integrations might give rise to a logarithm. The factors mc are applied in order to make the dimensions fit. It was proved that

$$\lim_{(m=0)} \tilde{W}^{(n)}=0.$$

This is only possible if $N \leqq 0$. This is equivalent to the fact that $\tilde{W}^{(n)}$ does not diverge stronger than

$$\tilde{W}^{(n)} \sim \left(\frac{e^2}{hc}\right)^n mc^2 \left(\lg \frac{P}{mc}\right)^n.$$

The Dispenser Cathode

A New Type of Thermionic Cathode for Gaseous Discharge Tubes

A. W. HULL
Research Laboratory, General Electric Company, Schenectady, New York
(Received May 13, 1939)

Modern oxide-coated cathodes have the characteristic that the initial high electron emission decreases gradually with time of operation. This characteristic is avoided by using initially-uncoated cathode members to which a coating material is dispensed at uniform rate from a "dispenser" located within the common heat-shield. Such cathodes, after initial aging, show an electron emission equal to the initial emission of oxide-coated cathodes, which does not decrease with time. In the cathodes which have been tested, the dispenser was a closely woven "stocking" of fine molybdenum wires, filled with granules of fused $BaO-Al_2O_3$ eutectic. It was maintained at 1150° C by current through it, and served as a radiation heater for the cathode. The electron-emitting members were clean molybdenum. These cathodes show undiminished electron emission at the end of three years life test with no indication of exhaustion.

1. INTRODUCTION

MODERN oxide-coated cathodes consist of nickel or platinum-nickel base to which the coating of barium and strontium oxides is applied once-for-all by means of a spray or dip. They represent the most efficient and most durable type of cathode in use today. Yet they have certain unfavorable features for heavy current tubes (above 25 amperes). The coating tends to flake off when applied to large areas, especially in the presence of mercury vapor. Its high electric resistance causes it to be heated by the emission current;[1] or to be electrolyzed, with liberation of oxygen.[2] It must be "activated," by chemical reaction with reducing elements in the base metal,[3] or with carbon deposited from the vapor phase,[4] or by electrolysis. It gradually evaporates, limiting the life of the cathode.[5] Long life has been attained under certain conditions,[6] though much remains still to be learned regarding the maintenance of emission during life.

2. STRUCTURE OF DISPENSER CATHODE

The dispenser cathode is similar to the heat-shielded oxide-coated cathode,[6] except in the method of applying the coating; the coating material is dispensed to the cathode members gradually and continuously throughout life, instead of being applied all at once. Fig. 1 shows the structure of a typical dispenser cathode. It differs from the oxide-coated cathode in two respects: (1) The electron-emitting members, in this case radial vanes, are clean molybdenum, instead of oxide-coated nickel. (2) The tungsten filamentary heater is replaced by a porous tubular filament, filled with activating material, which serves both as heater and as "dispenser" of coating.

The tubular filament may be a closely-wound helix of tungsten or molybdenum wire, or a tube of perforated molybdenum foil. But the most convenient form is a closely-woven cylindrical mesh or "stocking" of fine molybdenum wires, as shown in Fig. 1. It is easy to choose the dimensions of the stocking so that its voltage and current at the desired temperature are the same as those of the corresponding tungsten heater, thus making the dispenser cathodes interchangeable with oxide-coated cathodes. For example, a closely-woven stocking of 3-mil (0.075 mm) molybdenum wire, 2 mm in diameter and 20 cm long, requires 20 amperes at 5 volts to maintain its temperature at 1150°C in an enclosure at 850°C. Fig. 2 shows a cathode with such a dispenser, mounted in a tube of standard form.

[1] W. Espe, Zeits. f. tech. Physik **10**, 493 (1929).
[2] J. A. Becker, Phys. Rev. **34**, 1323 (1929).
[3] M. Benjamin, Phil. Mag. **20**, 1 (1935).
[4] C. H. Prescott, Jr. and J. Morrison, J. Am. Chem. Soc. **60**, 3047 (1938).
[5] M. Benjamin and H. P. Rooksby, Phil. Mag. **15**, 810 (1933); Claassen and Veenemans, Zeits. f. Physik **80**, 342 (1933).
[6] A. W. Hull, Trans. A. I. E. E. **47**, 753 (1928).

As "activating" filler for the dispenser we have used barium oxide in the form of small granules of fused Barya-alumina eutectic (70 BaO, 30 Al_2O_3 by weight). In this eutectic the free barium oxide is so well protected that the granules can be kept in air for years, without deterioration. On the other hand, the rate of diffusion of BaO at 1150°C is sufficient to build up nearly equilibrium pressure of BaO inside the stocking, thus insuring a constant rate of supply of BaO, and of its reaction products with Mo, to the cathode members, as long as any barium oxide remains inside the stocking. This constancy of dispensing rate makes possible the constant electron emission throughout life, which is characteristic of dispenser cathodes.

The optimum value of porosity of the dispenser is that which supplies active material at just the right rate when the dispenser is at the required temperature, *viz.*, the temperature at which it reacts with BaO to give the required active product. In the case of Mo this temperature is between 1150° and 1200°C. The desired rate of dispensing is the rate just equal to that at which the coating is lost by evaporation from the cathode surface. The chemical tests given in Table VI show an accumulation of BaO on the cathode surfaces, indicating that the dispensing rate was higher than necessary. The porosity in this case was approximately 0.001, as evidenced by the ratio of the observed evaporation to the calculated value for a free surface.

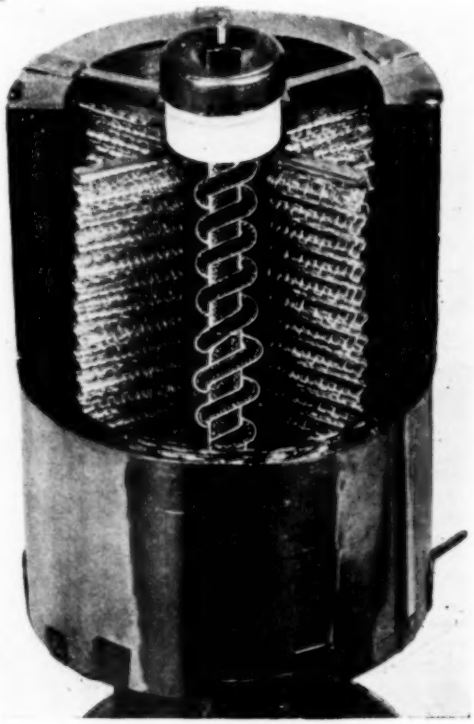


FIG. 1. Dispenser cathode, showing structure. The filamentary tungsten heater is replaced by a tubular heater, containing BaO, which continually dispenses coating to the vanes.

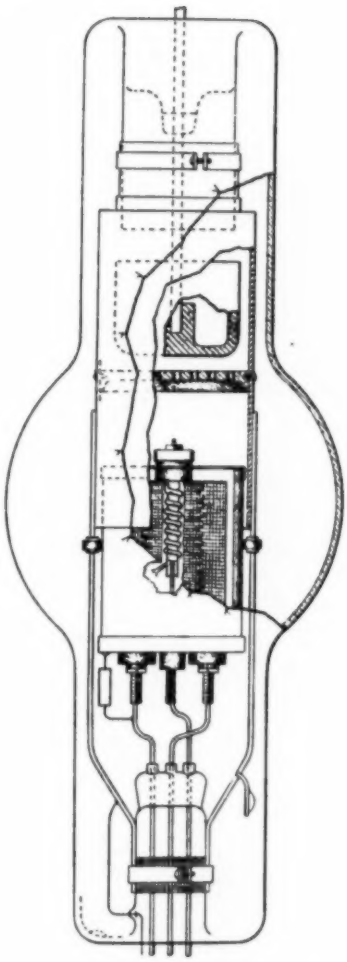


FIG. 2. Thyratron embodying dispenser cathode of Fig. 1.

Figure 3 shows an alternative form of cathode, in which the same stocking is compressed axially to form a straight tube, which is mounted axially in the cathode. The volt-ampere characteristics of a given stocking are approximately the same when thus compressed as in the extended form of Fig. 1. An experimental tube embodying this cathode is shown in Fig. 4.

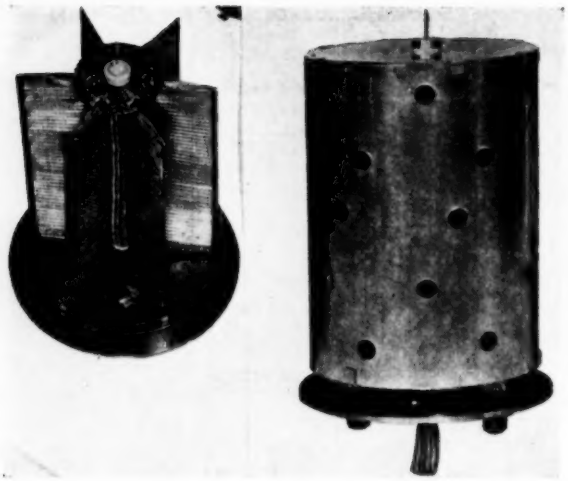


FIG. 3. Cathode embodying alternative form of dispenser. The molybdenum "stocking" is compressed to form a straight tube.

3. ELECTRON EMISSION

A number of tubes of the form shown in Fig. 2 have been built and tested. The electron-emitting vanes are molybdenum foil, 3 mils (0.075 mm) thick, to which molybdenum gauze has been welded on both sides, to increase the surface, giving a total surface area of 200 cm². The heat shield is insulated and plays no part in the emission; but it serves the important function of distributing the coating material, by re-evaporation, in addition to conserving heat and coating material.

Table I gives the electron emission of one of these tubes after 1000 hours of operation. The cathode in this tube was equipped with a platinum to platinum −10 percent rhodium thermocouple, whose junction was welded to the center of one of the vanes. Previous tests in a bell jar had shown that the average temperature thus determined represented all parts of the cathode, except the very lowest portions which were cooled by the supports, within about 10°C.

The method of determining the emission needs explanation: Since the currents are large enough to change the temperature of the vanes, it is necessary to make instantaneous observations. This is done by sending a single half-cycle of alternating current through the tube, and observing the voltage drop with a cathode-ray oscillograph. This drop has a "flat" constant value of about 15 volts for small currents; for larger currents it rises slightly in the middle of the cycle, representing the added drop necessary to extract electrons from the deeper portions of the cathode. When the emission limit is reached, an instantaneous demand for more emission causes the drop to increase abruptly. It is this limit, or a value just short of it, which is determined.

In Fig. 5 this limiting value of emission is shown recorded by a magnetic oscillograph, for the same tube at two temperatures; and these values are included in Table I and shown in Fig. 8. The method of taking these oscillograms was, first, to find the limit as described above, and adjust the circuit resistance; then, after waiting several minutes for the temperature to become constant, make the potentiometer reading of the thermocouple voltage, giving the temperature at the center of the vanes; then close the circuit and trip the oscillograph shutter in as close succession as possible. It was estimated that the oscillograph record began about 3 cycles after the current started, and that the maximum change in temperature before recording was 2°C.

The electron emission shown in Fig. 5 for the dispenser cathode can be judged most easily by comparison with oxide cathodes. Traces (c) and (d) of Fig. 5 show a similar test of an oxide cathode of standard type, which has a very high initial emission. This tube also was equipped with a thermocouple. The test was made with the tube in its initial condition, after the customary 24 hours aging. Trace (d) shows a current

TABLE I. *Electron emission tests of tubes similar to Fig. 2, with oxide cathode and dispenser cathode, respectively.*

TYPE OF CATHODE	CATHODE TEMP. °C	CATHODE WATTS	Hg PRESSURE, MM	ELECTRON EMISSION AMP.	AMP./CM²	AMP./WATT
Dispenser (data from oscillograms, Fig. 5)	714	60	0.014	45	0.225	0.75
	786	72	0.030	355	1.77	4.9
Dispenser (cathode-ray oscillograph tests)	734	63	0.021	107	0.53	1.70
	785	72	0.027	225	1.22	3.12
	831	80	0.023	360	1.80	4.50
Oxide (data from oscillograms, Fig. 5)	831	90	0.012	211	1.17	2.33

just less, trace (c) just greater, than the limit. The sudden drop in voltage at the middle of the cycle in trace (c) evidences the formation of a "cathode spot" on the oxide surface, indicating that the emission limit was exceeded.

Similar tests were made on the tube shown in Fig. 4, which was likewise equipped with a Pt-Pt Rh thermocouple, welded to the center of one vane, and carefully shielded. In this case an automatic timing device was employed to insure a duration of current of only one half-cycle. The data are given in Table II. The cathode was not yet fully activated, having been "aged" for a few hours with 40 amperes emission.

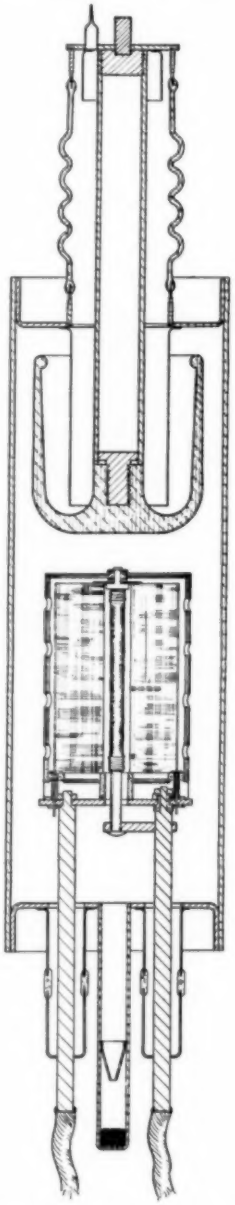


FIG. 4. Rectifier tube embodying cathode of Fig. 3.

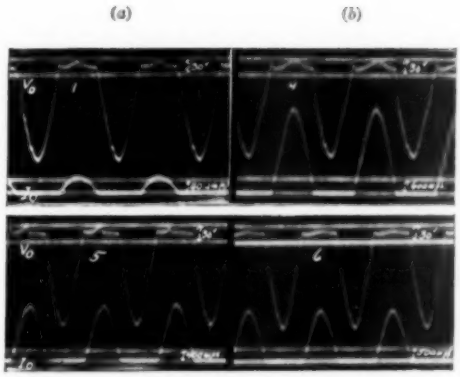


FIG. 5. Oscillograms of current and voltage drop of tubes similar to Fig. 2. The zero lines of the voltage and current traces are indicated by V_0 and I_0, respectively. Direct-current calibrations are shown at 30 volts and 40–60 amperes. (a) Dispenser-cathode tube, cathode temperature 714°C. (b) Same, cathode temperature 786°C. The current at middle of cycle is 355 amperes. (c) Oxide cathode tube, cathode temperature 835°C. The break at middle of cycle shows formation of cathode spot. (d) Same, slightly smaller current. Current at middle of cycle 211 amperes.

4. RICHARDSON CONSTANTS OF DISPENSER CATHODE

The practical forms of cathode described above are not suitable for determination of the electron-emission constants. The special cathode used for this purpose is shown in Fig. 6. The cathode proper is a massive molybdenum cylinder, 1.5 mm thick, 1.90 cm internal diameter and 1.50 cm long, to the outside of which is welded a Pt-Pt Rh thermojunction. Inside of this cylinder is mounted a molybdenum dispenser tube filled with $70BaO+30Al_2O_3$ granules, as described in Section 2. Radiation shields, at each end of the cylinder, were supported on the dispenser mount and served, with it, as anode for the electron currents. The tube was degassed, filled with neon at 2 mm pressure, and sealed off. Activation was accomplished by passing a current of 1 to 5 amperes at 15 volts from the cylinder, as cathode, to the dispenser and end plates, as anode, for a total of six hours.

Emission tests were made at 17 volts, in order to avoid ionization of the neon; hence the emission which was measured is essentially "field-free" emission. A space-charge test at one

milliampere showed six-percent change in current for 25-percent change in voltage. The maximum current used for the emission tests was 470 microamperes.

The results of the emission tests are given in Table III and Fig. 7. All the readings, which were taken on five different days, have been included, in order to show the remarkable constancy of the emission. It will be seen that only two readings deviate from the common straight line by more than the experimental error, ± 1°C, of the temperature measurements.

The data in Fig. 7 yield the constants, in the Richardson equation:

$$A = 0.85 \text{ amp./cm}^2/\text{deg}^2.$$
$$\varphi = 1.215 \text{ electron volts.}$$

For convenience of comparison, the emission represented by these constants in the range of practical operating temperatures has been plotted in Fig. 8, together with the best available data on oxide-coated cathodes.

5. Life

The most important feature of the dispenser cathode is its maintenance of electron emission during life. Life tests were made on small tubes of the form shown in Fig. 9. The results to date are given in Table IV. This table contains all the tubes of this type* that were tested, and it will be noted that all except two are still operating; one failed after 14,173 hours from an unrecorded cause, the other was broken by accident during test, after 23,850 hours.

Table II. *Electron emission of cathode shown in Fig. 4, after initial aging.*

Cathode Temp. °C	Cathode Watts	Hg Pressure mm	Electron Emission amp.	amp./cm²	amp./watt
691	300	0.020	120	0.080	0.40
738	365	0.022	240	0.160	0.61
788	450	0.030	520	0.350	1.26
838	545	0.033	730	0.49	1.34
845	560	0.011	760	0.51	1.36
888	660	0.035	1080	0.73	1.64
930	780	0.025	1400	0.94	1.80

* These tubes all contained mercury vapor. Some tubes containing neon were tested also, but the heating of the cathode by the neon discharge, and the tendency of the neon to clean up, make the results less definite.

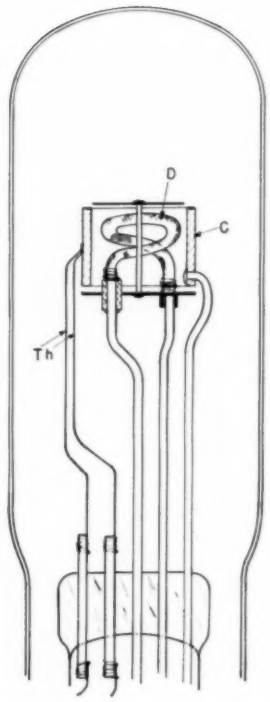


Fig. 6. Special dispenser cathode used for study of electron emission constants. C=cathode; D=dispenser; Th=thermocouple. The massive form is designed to give uniform temperature and enable use of thermocouple.

One of these tubes, No. 35879, was retested on April 12, 1939, after approximately 30,000 hours operation. The results are given in Table V, and are shown on curve 1 in Fig. 8. The emission tests were made by instantaneous observation of current and voltage drop, with the aid of the cathode-ray oscillograph, as already described in connection with the data of Table I. The cathode temperatures were determined by optical pyrometry as follows: Pyrometer measurements on three clean tubes, duplicates of the tube under test, gave identical temperatures within 10°C for the same watts. One of these was then carefully pyrometered, determining temperature as a function of watts. These readings, converted to true temperatures by correction for emissivity and glass absorption, fitted well on a log W *vs.* 4 log T straight line plot. This plot was used to convert watts to temperature for tube 35879, which was too black for observation. The results are believed to be trustworthy within 20°C. Correction has been made for the emission of the dispenser itself (connected to cathode in 35879), by determining the emission from the dispenser of a simi-

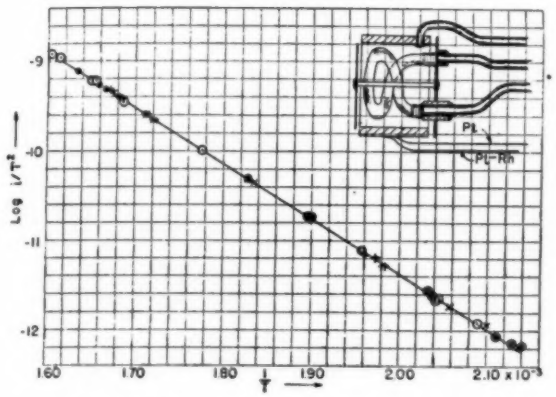


FIG. 7. Richardson plot of emission from dispenser cathode, of 20 cm² area, measured at 17 volts. The plot includes data taken on five different days (distinguished by different symbols). It is represented by the equation $i=0.85\ T^2e^{-1.215/kT}$ amp./cm² (k=Boltzmann constant in electron volts$=8.62\times10^{-5}$).

TABLE III. *Electron emission of dispenser cathode.*

Date	Thermocouple Reading Millivolt	Temp. °K	Electron Emission Microampere	1000/T	$\log_{10} i/T^2$
Mar. 27	2.67	612	285	1.634	0.881−10
	2.59	603	198	1.658	0.736−10
	2.56	600	170	1.667	0.674−10
	2.51	595	140	1.681	0.597−10
Mar. 28	2.42	584	90	1.712	0.421−10
	2.39	581	73	1.721	0.335−10
Apr. 3	2.50	593	127	1.686	0.558−10
	2.55	598	172	1.672	0.682−10
	2.54	597	166	1.675	0.668−10
	1.715	505	1.30	1.980	0.707−12
	1.70	504	1.295	1.984	0.707−12
	1.73	507	1.66	1.972	0.810−12
	1.75	510	1.79	1.961	0.838−12
	1.73	507	1.625	1.972	0.801−12
	1.58	490	0.524	2.041	0.339−12
	1.55	486	0.426	2.058	0.256−12
	1.485	476	0.262	2.101	0.063−12
	1.395	468	0.141	2.137	0.809−13
	2.065	545	12.95	1.835	0.641−11
	2.08	547	14.3	1.828	0.679−11
Apr. 4	1.834	467	0.151	2.141	0.840−13
	1.484	479	0.271	2.088	0.072−12
	1.333	461	0.095	2.169	0.650−13
	1.573	489	0.546	2.045	0.359−12
	1.90	527	5.14	1.898	0.267−11
	2.223	563	32.2	1.776	0.007−10
	2.62	606	220	1.650	0.777−10
	2.75	620	404	1.613	0.022−9
	2.79	624	470	1.603	0.082−9
	2.615	605	223	1.653	0.785−10
	2.495	593	123	1.686	0.544−10
Apr. 6	1.41	470	0.165	2.128	0.873−13
	1.44	474	0.194	2.110	0.936−13
	1.59	491	0.59	2.037	0.390−12
	1.60	492	0.64	2.033	0.422−12
	1.76	511	2.07	1.957	0.899−12
	1.90	527	4.8	1.898	0.239−11
	1.91	528	5.1	1.894	0.262−11
	2.08	547	15.3	1.828	0.708−11

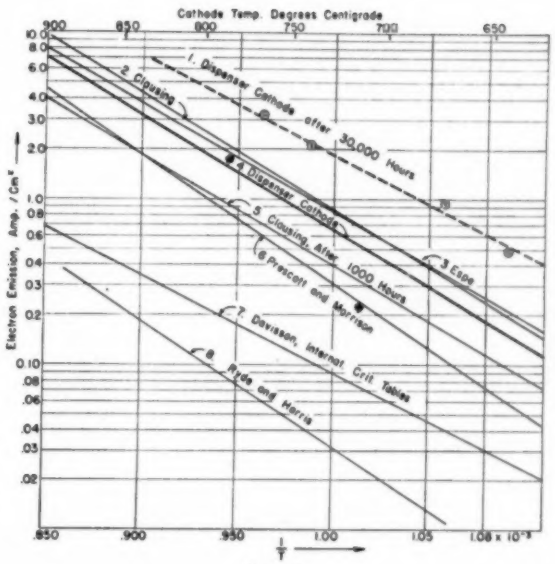


FIG. 8.

1. Dispenser-cathode, after 30,000 hours operation (See Table V).
2. P. Clausing;[7] "well activated" nickel filament, coated with BaO+SrO: $A=0.96$, $\varphi=1.20$.
3. W. Espe, reference 1; barium on oxidized tungsten: $A=0.3$, $\varphi=1.10$.
4. Dispenser-cathode (from data of Table III and Fig. 7): $A=0.85$, $\varphi=1.215$.
5. P. Clausing, same as 2, after 1000 hours operation: $A=0.25$, $\varphi=1.15$.
6. Prescott and Morrison, reference 4; Pt-Ni, coated with BaO+SrO and activated by C from CH_4: $A=2.45$, $\varphi=1.37$.
7. *International Critical Tables;* Pt-Ni, coated with BaO+SrO (data from Bell Laboratories): $A=10^{-2}$, $\varphi=1.0$.
8. Ryde and Harris;[8] Ba-O-W: $A=0.18$, $\varphi=1.34$.

lar tube with insulated dispenser, which had been operated for the same length of time. This correction, approximately three percent, was subtracted from the observed emission of tube 35879. The data in Table V and Fig. 8, thus corrected, represent the emission from the molybdenum cup.

This emission is quite high, and would still be high if one allowed for a possible error of 20° or even 30° in temperature. The measurements of current and temperature are not accurate enough to justify derivation of Richardson constants, but they justify the conclusion that the electron emission of this cathode, after 30,000 hours of life-test operation, is higher than any initial emission that has been observed.

* Dispenser-cathode after 1000 hours, from oscillogram, Fig. 5.

[7] P. Clausing, unpublished data, quoted by J. H. DeBoer, *Electron Emission and Absorption Phenomena* (MacMillan, 1935), p. 361.

[8] J. W. Ryde and N. L. Harris, unpublished data, quoted by A. L. Reimann, *Thermionic Emission* (Chapman and Hall, London, 1934), p. 165.

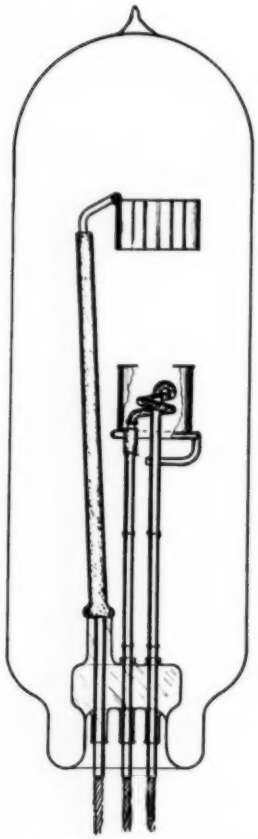


FIG. 9. Dispenser-cathode tube used for life tests. The dispenser is the same as in Fig. 1.

In applying the results of these life tests to practical cathodes for heavy current tubes like that shown in Fig. 1, one may proceed with considerable confidence. If the temperatures of dispenser and cathode enclosure are the same in the two cases, and also the ratio of dispenser area to cathode opening, then one may expect the same density and vapor pressure of active material, and hence the same electron emission per unit of cathode surface throughout life.

What is the limit of life of a dispenser cathode? The data in Table VI give some information on this subject. These data show that less than half of the barium oxide left the "stocking" in 20,000 hours, and half of this was still in the cathode enclosure. Only one of these tubes (No. 35168, broken during test) contained mercury vapor. This tube shows a net loss of BaO from the cathode enclosure of only 7 mg, or 2.4 percent, in 23,850 hours, which would indicate a possible life of many years. The neon tubes ran about 50° hotter, on account of the high ionization energy of neon, which accounts for their greater evaporation. However, the initial amount of activating material in the dispenser is not limited; hence a very long life appears possible under any reasonable operating conditions, by proper design. Perhaps the most important consideration, if these data are substantiated by further tests, will be the ability to predict life, from the amount of material in the dispenser and the rate of dispensing.

TABLE IV. *Life tests of barium-dispenser cathodes.*

Tube Number	Life Tests Conditions: Mercury Temperature deg. C	Cathode Watts	Anode Current Amperes	Hours Operation at 11/20/38	Electron Emission 11/20/38: at 770°, amp.	Hg 40°C amp./cm²
35036	<40	53	3.0 d.c.	31334+	22	1.1
34743	60	47	9.5 max. a.c.	32178+	12.6	0.63
35907	40–50	47	5.7 "	26279+	25	1.25
35867	40–50	47	5.7 "	26279+	25	1.25
35879	40–50	47	8.8 "	26279+	70	3.5
35933	40–50	47	9.5 "	14173	—	—
36593	30	45	9.5 "	19368+	39	1.95
35168	75	47	13.5 "	23850	*	*
35181	75	47	20.5 "	27980+	12.6	0.63
35174	72	47	7.9 "	27980+	25	1.25
35082	80	47	13.2 "	23850+	12.3	0.61
			Ave.	25414+	27.0	1.35

* Broken during test; used for chemical analysis.

TABLE V. *Electron emission of experimental dispenser cathode after 30,000 hours operation.*

Cathode Temp. °C	Cathode Watts	Hg Pressure mm	Electron Emission: amp.	amp./cm²
597	18.8	0.012	2.0	0.10
644	25.5	0.012	10.0	0.50
672	30.3	0.0155	18.9	0.95
740	44.0	0.010	43.1	2.15
766	50.5	0.0165	64.2	3.21

TABLE VI. *Distribution of barium oxide in dispenser-cathode tubes.*

Tube Number	Hours of Operation	Gas	Milligrams of BaO: Stocking	Cup	Bulb	Anode	Total
35168	23,850	Hg	242.5	43.8	7.0	0.0	293.3
34932	19,044	Ne	135.3	50.3	4.4	0.0	190.0
34949	21,644	Ne	97.8	43.8	32.7	12.8	187.1
34977	20,181	Ne	103.3	63	37.4	16.8	220.6
blank	0*	Ne	191.5	—	—	—	191.5

* New tube.

In contemplating long operating life the question of vacuum deserves consideration. Some observations were made last year on power Thyratrons which had been in operation for two years at constant current (175 amperes, 120 degree conduction). These tubes were tested in the laboratory for ability to operate at high voltage, and the results compared to initial tests on the same tubes. In every case the voltage tolerance was found to be higher than at the start. This shows that vacuum is not impaired by operation under these conditions.

JULY 1, 1939 PHYSICAL REVIEW VOLUME 5

Change of Resistance in a Magnetic Field

LEVERETT DAVIS, JR.
California Institute of Technology, Pasadena, California
(Received April 18, 1939)

The general theory of the Hall effect and the change of resistance in a magnetic field expresses these quantities in terms of a number of integrals over the surface of the Fermi distribution. The values of these integrals depend upon the form of the electron energy and the relaxation time as functions of the wave vector. If the free electron situation is assumed, the Hall effect has the right order of magnitude, but there is no change of resistance. This can be seen from a qualitative consideration of the effect of the fields on the distribution function. A general form for the functions in other cases can be obtained as an expansion in spherical harmonics with the symmetry of the crystal lattice. The results can then be expressed in terms of the coefficients in this expansion. When only the first two harmonics are retained, the computed change of resistance and Hall effect are close to the observed values. However, contrary to the available observations, the ratio of the transverse to the longitudinal change of resistance shows a minimum value of about four. It seems improbable that this result could be changed in any material way by the inclusion of higher series members, so that if the experimental results are to be taken as reliable, doubt is thrown on the general method of treatment.

§1. INTRODUCTION

CHANGE of resistance in a magnetic field cannot be interpreted in terms of a free electron picture. But since such a change is observed in the alkalies, it is of interest to see if it can be understood as a small departure from the free electron situation. Jones and Zener[1] have given a theory for this effect that seems to give quite satisfactory numerical results in the case of lithium. However, their approximate method of evaluating the integrals made it necessary to apply, later, an estimated correction factor of about six. This paper attempts to give a method of evaluating the integrals that will enable one to get analytically as good an approximation as is desired, provided the surfaces of constant energy are not too irregular. Also it is not assumed that the relaxation time is a function of the energy only. §2 gives a summary of the general theory, in Wilson's[2] notation, and §3 gives a physical picture of the situation, in which particular attention is given to the free electron case. In §4 the energy and the relaxation time are expressed in terms of series of cubically symmetric spherical harmonics. By a proper choice of independent variables, the integrals can all be evaluated in terms of the coefficients of the spherical harmonics. If the first two terms of the series are taken as a satisfactory approximation, expressions are obtained for the conductivity, the Hall coefficient, and the two coefficients, B_t and B_l, of the change of resistance in a magnetic field. In §5 the conclusions that may be drawn from these expressions are discussed. It is found that theory and experiment give radically different values of B_t/B_l. Experimentally this ratio is about unity, while no choice of parameters can give a theoretical value less than 4. The expressions for B_t and B_l show that the variation of the

[1] H. Jones and C. Zener, Proc. Roy. Soc. **A145**, 268 (1934).

[2] A. H. Wilson, *The Theory of Metals* (Cambridge Univ. Press, 1937), Chapter V.

relaxation time with direction, as well as with energy, is nearly as important in all these effects as the fact that the surfaces of constant energy are not spheres.

§2. General Results

We will specify the state of an electron in the usual manner by the wave vector $\mathbf{k}$ whose Cartesian components are (k_1, k_2, k_3). The energy eigenfunctions of the electrons when unperturbed by external fields can be taken as

$$\psi_{\mathbf{k}}(\mathbf{r}) = \exp(i\mathbf{k}\cdot\mathbf{r})u_{\mathbf{k}}(\mathbf{r}),$$

where $u_{\mathbf{k}}(\mathbf{r})$ has the periodicity of the lattice. The number of electrons per unit volume whose wave vectors lie in the range (dk_1, dk_2, dk_3) is

$$(1/4\pi^3)f(\mathbf{k})dk_1dk_2dk_3.$$

We omit any dependence on $\mathbf{r}$ since we are going to be interested only in the case in which the temperature and composition of the metal are independent of $\mathbf{r}$.

If there exist in the metal an electric field, $\boldsymbol{E}$, and a magnetic field, $\mathbf{H}$, then the condition that the distribution function, $f(\mathbf{k})$, remains constant in time under the influence of the fields and the collisions of the electrons with the lattice points is given by Boltzmann's equation,

$$-(\epsilon/\hbar)[\boldsymbol{E}+\mathbf{v}_{\mathbf{k}}\times\mathbf{H}/c]\cdot\text{grad.}_{\mathbf{k}} f + (f-f_0)/\tau(\mathbf{k}) = 0, \quad (1)$$

where $-\epsilon$ is the charge on an electron, $\mathbf{v}_{\mathbf{k}}$ is the velocity associated with the state $\mathbf{k}$, and f_0 is the value of f when $\boldsymbol{E}=\mathbf{H}=0$. We have assumed that a relaxation time, $\tau(\mathbf{k})$, can be defined so that we can use this form of Boltzmann's equation. A relaxation time can be used if the temperature is greater than the Debye characteristic temperature, and perhaps under other conditions. Eq. (1) is only valid if $\epsilon H\tau/2mc \ll 1$.

To solve Eq. (1), take the direction of the magnetic field as the z axis and substitute for $\mathbf{v}$ its value in terms of $\mathbf{k}$,

$$\mathbf{v}_{\mathbf{k}} = (1/\hbar)\,\text{grad.}_{\mathbf{k}} E(\mathbf{k}), \quad (2)$$

where $E(\mathbf{k})$ is the energy associated with the state $\mathbf{k}$. It should be noted that E refers to the energy of an electron while $\mathbf{E}$ always refers to the electric field.

Let

$$f = f_0 - \Phi(\mathbf{k})\partial f_0/\partial E.$$

The equation which $\Phi(\mathbf{k})$ is to satisfy is

$$\Phi(\mathbf{k})/\tau(\mathbf{k}) + (\epsilon/\hbar)\boldsymbol{E}\cdot\text{grad.}_{\mathbf{k}} E - (\epsilon H/\hbar^2 c)\Omega\Phi(\mathbf{k}) = 0, \quad (3)$$

where the product of $\boldsymbol{E}$ and Φ is neglected, and Ω is the operator

$$\Omega = \frac{\partial E}{\partial k_2}\frac{\partial}{\partial k_1} - \frac{\partial E}{\partial k_1}\frac{\partial}{\partial k_2}.$$

This process is equivalent to a development of f as a power series in the components of $\boldsymbol{E}$ and the neglect of all terms of higher than the first degree.

To solve Eq. (3) we write $\Phi(\mathbf{k})$ as a power series in H, and obtain

$$\Phi(\mathbf{k}) = -(\epsilon/\hbar)\{\tau\boldsymbol{E}\cdot\text{grad. } E + (\epsilon H/\hbar^2 c)\tau\Omega(\tau\boldsymbol{E}\cdot\text{grad. } E) + (\epsilon H/\hbar^2 c)^2\tau\Omega[\tau\Omega(\tau\boldsymbol{E}\cdot\text{grad. } E)] + \cdots\}. \quad (4)$$

This is one particular integral of (3) and may or may not be the solution we desire. The general solution of (3) is obtained by adding to (4) the general solution of

$$\Phi(\mathbf{k}) = \tau(\mathbf{k})(\epsilon H/\hbar^2 c)\Omega\Phi(\mathbf{k}) = (\tau\epsilon H/\hbar^2 c)[\text{grad.}_{\mathbf{k}}\Phi(\mathbf{k})]\cdot[\text{grad.}_{\mathbf{k}} E(\mathbf{k})]\times\mathbf{e}_3,$$

where $\mathbf{e}_3$ is a unit vector in the k_3 direction. Consider the curve of intersection of any surface of constant energy with the plane $k_3=0$. The vector $(\text{grad.}_{\mathbf{k}} E)\times\mathbf{e}_3$ will be directed along this curve. In all cases of interest $E(\mathbf{k})$ is such that this vector is directed in the clockwise sense of describing the curve. Since $\tau(\mathbf{k})$ is essentially positive, our equation shows that the component of $\text{grad.}_{\mathbf{k}}\Phi$ in the direction of $(\text{grad.}_{\mathbf{k}} E)\times\mathbf{e}_3$ is always of the same sign as Φ. Hence if we traverse our curve in the clockwise sense, $|\Phi|$ will increase unless $\Phi=0$. Therefore the only single-valued solution of the homogeneous equation is $\Phi\equiv 0$. Hence (4) is the desired solution.

The reason that we can terminate the series in the components of $\boldsymbol{E}$ with the linear terms while we must include at least the quadratic term in the series in H is that the magnetic force is very much larger than the electrostatic for electrons

whose representative points are at the surface of the Fermi distribution. If the current density in silver is 10 amp. cm^{-2}, and if the magnetic field is 1000 gauss, the magnetic force is 10^8 times as great as the electric force.

We can get expressions for the conductivity, the Hall coefficient, and the change of resistance in a magnetic field by substituting the value of Φ found above in the equation giving the current density,

$$\mathbf{J}=-(\epsilon/4\pi^3)\int \mathbf{v}_k f(\mathbf{k})dV$$
$$=(\epsilon/4\pi^3\hbar)\int \text{grad.}_k E(\partial f_0/\partial E)\Phi dV,$$

where dV is an element of volume in $\mathbf{k}$-space.

To do this we introduce the following abbreviations:

$$I_1=-\int \tau(\partial f_0/\partial E)(\partial E/\partial k_1)^2 dV,$$

$$I_2=-\int \tau(\partial f_0/\partial E)(\partial E/\partial k_2)^2 dV,$$

$$I_3=-\int \tau(\partial f_0/\partial E)(\partial E/\partial k_3)^2 dV,$$

$$I_4=I_2I_6-I_7^2,$$

$$I_5=\int \tau\frac{\partial f_0}{\partial E}\frac{\partial E}{\partial k_3}\Omega\left\{\tau\Omega\left(\tau\frac{\partial E}{\partial k_3}\right)\right\}dV$$
$$=-\int \tau\frac{\partial f_0}{\partial E}\left\{\Omega\left(\tau\frac{\partial E}{\partial k_3}\right)\right\}^2 dV, \quad (5)$$

$$I_6=\int \tau\frac{\partial f_0}{\partial E}\frac{\partial E}{\partial k_1}\Omega\left\{\tau\Omega\left(\tau\frac{\partial E}{\partial k_1}\right)\right\}dV$$
$$=-\int \tau\frac{\partial f_0}{\partial E}\left\{\Omega\left(\tau\frac{\partial E}{\partial k_1}\right)\right\}^2 dV,$$

$$I_7=\int \tau\frac{\partial f_0}{\partial E}\frac{\partial E}{\partial k_1}\Omega\left(\tau\frac{\partial E}{\partial k_2}\right)dV$$
$$=-\int \tau\frac{\partial f_0}{\partial E}\frac{\partial E}{\partial k_2}\Omega\left(\tau\frac{\partial E}{\partial k_1}\right)dV.$$

The second form of each integral is obtained by integrating the first form by parts.

For those metals in which $E(\mathbf{k})$ and $\tau(\mathbf{k})$ are even functions of k_1, of k_2, and of k_3, we find that the electrical conductivity in the absence of a magnetic field is

$$\sigma=J_x/\boldsymbol{E}_x=(\epsilon^2/\pi\hbar^2)I_1, \quad (6)$$

that the Hall coefficient is

$$R=\boldsymbol{E}_y/HJ_x=\boldsymbol{E}_y/H\sigma\boldsymbol{E}_x=-(4\pi^3/c\epsilon)I_7/I_1I_2, \quad (7)$$

and that the coefficients of the change of resistance in magnetic fields that are perpendicular and parallel, respectively, to the direction of the current are

$$B_t=(\sigma_0-\sigma)/\sigma H^2=(\epsilon/\hbar^2c)^2I_4/I_1I_2 \quad (8)$$

and

$$B_l=(\sigma_0-\sigma)/\sigma H^2=(\epsilon/\hbar^2c)^2I_5/I_3. \quad (9)$$

By Schwarz's inequality the change of resistance is always an increase. These formulas are equivalent to those given by Wilson.[2]

§3. Physical Picture

A careful study of the above equations enables one to get a physical picture of what is going on. In the absence of external forces, the distribution of representative points in $\mathbf{k}$-space is given by the Fermi function, $f_0(\mathbf{k})$. The density of distribution depends on the energy only. The application of a force, $\mathbf{F}$, causes the distribution of representative points to drift in the direction of the force, since rate of change of the state of an electron is given by

$$d\mathbf{k}/dt=\mathbf{F}/\hbar.$$

When we assume the Boltzmann equation in the form of Eq. (1), we are really assuming that the effect of the collisions of the electrons with the ions is such that the actual distribution, f, differs at each point from the normal distribution, f_0, by an amount that is proportional to the rate at which f would tend to be changed by the drift produced by the force. We are not assuming that a representative point drifts for a certain time or distance, on the average, before a collision changes its state.

Our series solution of Boltzmann's equation is really a solution by successive approximations Rather than consider the effect of the drift

produced by the external forces on the final unknown distribution, we first calculate the change in distribution on the assumption that only the representative points of the undisturbed function, f_0, drift. Hence from Eq. (1) we see that in the first approximation we must add to the equilibrium distribution function the term

$$-\tau\mathbf{F}\cdot\text{grad.}_{\mathbf{k}}\, f_0.$$

We get the second approximation by considering the drift of those points given by the first approximation and obtain

$$-\tau\mathbf{F}\cdot\text{grad.}_{\mathbf{k}}[-\tau\mathbf{F}\cdot\text{grad.}_{\mathbf{k}}\, f_0].$$

Each additional term is found by considering the action of the forces on the representative points given by the previous term.

Since the forces on conduction electrons due to magnetic fields in metals are, in general, much greater than those due to electric fields, we include the electric forces in the first term only of this series of approximations. We do not include the magnetic forces in this first term since the drift produced by them is perpendicular to $\mathbf{v}=\text{grad.}_{\mathbf{k}}\, E/\hbar$. It is, therefore, along the surfaces of constant energy and produces no change in the undisturbed distribution. The first approximation gives the ordinary conductivity, the second gives the Hall effect, and the third gives the change of resistance in a magnetic field.

Let us consider our picture of the free electron case. We may assume complete degeneracy since Sommerfeld and Frank[3] have shown that the departure from complete degeneracy gives only about one-ten-thousandth of the observed change of resistance in a magnetic field. Therefore when no external field is applied, the density of representative points is uniform inside a certain sphere and zero outside. When an external field is applied, the change in the distribution function is given by a surface density of representative points on this sphere. We can show this in Fig. 1 by taking a cross section in $\mathbf{k}$-space through the plane $k_3=0$. The circle represents the interaction of the sphere with this plane and the width of the surrounding band indicates the surface density of representative

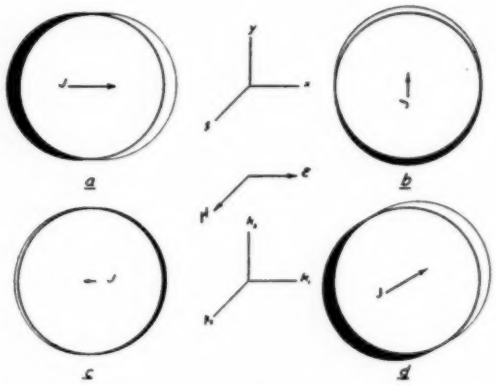


FIG. 1. The free electron case. Since the charge of the electron is taken to be negative, the representative points drift in a direction opposite to that of the electric field.

points. Black indicates a positive and white a negative surface density. For the case in which the electric field is in the x direction and the magnetic field is in the z direction, Fig. 1 (a) shows the effect of the electric field on the undisturbed distribution. Fig. 1 (b) gives the second approximation and shows the effect produced by the action of the magnetic field on the surface density of representative points shown in (a). Fig. 1 (c) gives the third approximation and shows the effect produced by the magnetic field on the points that make up the second approximation. The actual distribution, shown in (d), is obtained by superposing these three terms.

It is evident that if the electric field is in the x direction, then the current is not. But all the effects are defined for the case in which the current is in the same direction, with and without the magnetic field. Hence the Hall effect arises from the fact that to eliminate the current indicated in (b) we will need to apply an electric field in the negative y direction. Fig. 1 (c) would seem to indicate that the current is less when we have a magnetic field, that is, that the magnetic field increases the resistance. But if we consider the Hall effect of the electric field introduced to suppress the y component of the current, we see that if we have spherical symmetry, this effect completely cancels out and there is no change of resistance.

If the magnetic field is parallel to the current, the application of an electric field yields, if we have spherical symmetry, a first approximation that has cylindrical symmetry about an axis

[3] A. Sommerfeld and N. H. Frank, Rev. Mod. Phys. **3**, 1 (1931).

parallel to the magnetic field. Therefore the magnetic field produces no change in the distribution function and hence no change of resistance.

If we do not have spherical symmetry, we must remember that the velocity is in a direction perpendicular to the surfaces of constant energy. The interaction of effects is very complicated, but, by a sufficient extension of this kind of argument, it can be seen in a qualitative way that the effect of all departures from spherical symmetry is a tendency to increase the resistance in a magnetic field. If the departures are small, the Hall effect is not greatly changed. If the departures from spherical symmetry are large, then by proper choice of the shapes of the surfaces of constant energy it is easy to explain the fact that the Hall coefficient is positive in some metals.

§4. Calculations

In the case of the monovalent metals we expect that the surface of the Fermi distribution will lie entirely within the first Brillouin zone and will be nearly spherical. By expressing the dependent variables in terms of spherical harmonics, we can evaluate the integrals of §2 and get simple expressions for the various coefficients.

Our integrals are all of the form

$$I=\int(\partial f_0/\partial E)F(\mathbf{k})dV,$$

where $dV=k^2\sin\theta dk d\theta d\varphi$ is an element of volume in $\mathbf{k}$-space. If we use as our independent variables E, θ and φ, this becomes

$$I=\int_0^{2\pi}\int_0^{\pi}\int_0^{\infty}\frac{\partial f_0}{\partial E}F(\mathbf{k})k^2\sin\theta\left(\frac{\partial k}{\partial E}\right)_{\theta,\varphi}dEd\theta d\varphi.$$

We can do this since E is a monotonic function of k in the cases in which we are interested. Near the surface of the Fermi distribution all the factors in the integrand vary slowly compared to $\partial f_0/\partial E$, while elsewhere $\partial f_0/\partial E$ is practically zero. Hence

$$I=-\int_0^{2\pi}\int_0^{\pi}\left[\frac{\partial k}{\partial E}F(\mathbf{k})k^2\right]_{E=E_0}\sin\theta d\theta d\varphi, \quad (10)$$

where E_0 is the energy at the surface of the Fermi distribution.

The partial derivatives that we find in our integrands can be changed to the new variables by means of the following formulas from the calculus:

$$\begin{aligned}Ek_1&=(k\sin^2\theta\cos\varphi-k_\theta\sin\theta\cos\theta\cos\varphi\\&\qquad+k_\varphi\sin\varphi)(kk_E\sin\theta)^{-1},\\Ek_2&=(k\sin^2\theta\sin\varphi-k_\theta\sin\theta\cos\theta\sin\varphi\\&\qquad-k_\varphi\cos\varphi)(kk_E\sin\theta)^{-1},\\Ek_3&=(k\cos\theta+k_\theta\sin\theta)(kk_E)^{-1},\end{aligned}$$

where Ek_1 means the partial of E with respect to k_1, keeping k_2 and k_3 constant; while k_E means the partial of k with respect to E, keeping θ and φ constant.

Now the lattices of all the monovalent metals have cubic symmetry. Consequently when we expand $k(E,\theta,\varphi)$ and $\tau(E,\theta,\varphi)$ in series of surface harmonics, we need include only those surface harmonics that have cubic symmetry. The first two such surface harmonics are

$$Y_0^c=1$$

and $Y_4^c=P_4(\cos\theta)+\cos 4\varphi P_4^4(\cos\theta)/168$;

the next is of the sixth degree. P_n^n is Ferrers' associated Legendre function. Neglecting terms of the sixth degree and higher in our expansion, we write

$$\begin{aligned}k&=\alpha_0(E)+\alpha_1(E)Y_4^c,\\\tau&=\tau_0(E)+\tau_1(E)Y_4^c.\end{aligned}\quad (11)$$

Let us now transform the integrals of Eqs. (5) by the use of Eqs. (10) and (11). If we expand $[\alpha_0(E)+\alpha_1(E)Y_4^c]^n$, $[\alpha_0'(E)+\alpha_1'(E)Y_4^c]^n$, and $[\tau_0(E)+\tau_1(E)Y_4^c]^n$ by the binomial expansion, all the integrations can actually be carried out by means of formulas given by Gaunt.[4] If we write

$$a=\alpha_0(E_0),\quad a'=\left.\frac{d\alpha_0(E)}{dE}\right|_{E=E_0}\quad t=\tau_0(E_0),$$

$$A=\alpha_1(E_0)/\alpha_0(E_0),\quad B=\frac{1}{a'}\left.\frac{d\alpha_1(E)}{dE}\right|_{E=E_0},$$

$$T=\tau_1(E_0)/\tau_0(E_0),$$

[4] J. A. Gaunt, Trans. Roy. Soc. **228**, 151 (1929), Appendix.

and neglect the cubes of A, B and T compared to unity, we get

$$I_1=I_2=I_3=(4\pi/3)(t/a')a^2\{1+(4/21)\times[21A^2+2A(T-B)-B(T-B)]\},$$

$$I_4=(4\pi/3)^2(4/77)(t/a')^4a^2\times\{597A^2+46A(T-B)+37(T-B)^2\},$$

$$I_5=(4\pi/3)(80/231)(t/a')^3\{3A+T-B\}^2, \quad (12)$$

$$I_6=(4\pi/3)(t/a')^3\times\{1+(4/231)[2220A^2+600A(T-B)-33B(T-B)+532(T-B)^2]\},$$

$$I_7=(4\pi/3)(t/a')^2a\times\{1+(4/21)[30A^2+22A(T-B)-2B(T-B)+(T-B)^2]\}.$$

Substitution of these values in Eqs. (6)–(9) gives

$$\sigma=(4\epsilon^2/3h^2)(t/a')a^2\{1+(4/21)\times[21A^2+2A(T-B)-B(T-B)^2]\}, \quad (13)$$

$$R=-(3\pi^2/c\epsilon a^3)\{1+(4/21)\times[-12A^2+18A(T-B)+(T-B)^2]\}, \quad (14)$$

$$B_t=(4/77)(3\pi^2\sigma/c\epsilon a^3)^2\times\{597A^2+46A(T-B)+37(T-B)^2\}, \quad (15)$$

$$B_l=(80/231)(3\pi^2\sigma/c\epsilon a^3)^2\{3A+T-B\}^2. \quad (16)$$

We have eliminated t from the expressions for B_t and B_l by introducing σ from Eq. (13). We are to regard σ as being determined by experiment.

§5. Discussion of Results

In the evaluation of these integrals we have made three approximations. In the integration over E we have, as usual, assumed complete degeneracy. We have stopped at the second term in our expansions in surface harmonics and at the square terms in our binomial expansions of k, k' and τ. If it seemed desirable, more terms could be included in each expansion, but it does not seem likely that the general results would be changed much.

Reasonable values of the parameters, based on Jones and Zener's[1] work on lithium, give the correct orders of magnitude when inserted in the expressions (13)–(16). The consideration of the departures from spherical symmetry have a negligible influence on σ and R, but provide the essential part of B_t and B_l.

Since t and T always enter in the combinations t/a' and $(T-B)$, we get nothing by putting in a relaxation time that varies with direction that we could not obtain by suitable choice of the shape of the surfaces of constant energy. However, the extra parameters would give a means of getting better agreement between theory and experiment if the shape of the energy surfaces were given by other considerations.

If we consider the value of B_t/B_l given by Eqs. (15) and (16), we find that this ratio cannot be less than the value, 4.08, obtained when $A/(T-B)=1/6$, and that if $A/(T-B)$ does not lie between -0.07 and 5.5, B_t/B_l will be larger than 9. This disagrees with Jones and Zener's theoretical results and with the experimental results; in each the ratios are approximately one. There seem to be two ways of explaining this. One can say that the experimental results are not to be trusted, or one can say that some point has been overlooked in the development of the theory of magnetic effects in metals.

The author wishes to express his gratitude to Professor W. V. Houston for his continued interest in the research on which this paper is based, and for his very helpful advice at all stages of the work.

Diffraction Theory of Electromagnetic Waves

J. A. STRATTON AND L. J. CHU
Massachusetts Institute of Technology, Cambridge, Massachusetts

(Received May 15, 1939)

It has been shown by Larmor, Kottler and others that the classical method of calculating diffraction from the Kirchhoff formula in terms of a scalar light function cannot be applied directly to the electromagnetic field since it takes into account neither the vector character of the field nor the effect of charges along the contour of the opening. The field equations are integrated directly by means of a vector analog of Green's theorem. The results are applied to the calculation of diffraction of electromagnetic waves from a rectangular slit in a screen of infinite conductivity. The results are compared with an exact solution obtained recently by Morse and Rubenstein.

1. INTRODUCTION

RECENT advances in the technique of generating ultra-high frequency radio waves have stimulated interest in a number of problems of electromagnetic theory which heretofore have had only academic importance. A natural consequence of this trend towards short waves is an application of the methods of physical optics to determine the intensity and distribution of radiation from hollow tubes, horns, or small openings in cavity resonators. Now it is well known that the application of the Kirchhoff diffraction formula to an opening in an opaque screen involves several fundamental errors of principle. Nevertheless a remarkably good agreement between measured and calculated intensities is obtained in the region directly in front of the opening provided the wave-length is small compared to the size of the opening. If, on the other hand, the wave-length is relatively large, the radiation is distributed over a wide angle and the Kirchhoff formula proves to be definitely inaccurate. It must then be extended to account first for the vector character of the wave, and secondly for the discontinuities introduced about the contour of the opening. Both of these factors as sources of error have been recognized by previous writers. In the present paper we shall recall first the restrictions on the Kirchhoff formula when applied to a scalar wave function. A vector analog of Green's theorem will then be derived, from which an integral representation of electromagnetic fields can be obtained in a very simple manner. Next the theory is extended to surfaces over which the field vectors are discontinuous, and the results are finally compared to those obtained from the solution of a boundary value problem by rigorous methods.

2. THE KIRCHHOFF FORMULA FOR SCALAR WAVE FUNCTIONS

Let $\varphi(x, y, z)$ be any solution of

$$\nabla^2\varphi+k^2\varphi=0, \tag{1}$$

which is continuous and has continuous first derivatives throughout a closed domain V, and let φ_S, $(\partial\varphi/\partial n)_S$ denote the values of φ and its normal derivative on the bounding surface S. Then if the direction of the normal is outward from S, the integral

$$u(x', y', z') = \frac{1}{4\pi}\int_S\left[\left(\frac{\partial\varphi}{\partial n}\right)_S\frac{e^{ikr}}{r}-\varphi_S\frac{\partial}{\partial n}\left(\frac{e^{ikr}}{r}\right)\right]da \tag{2}$$

represents a discontinuous function which at all interior points x', y', z' is equal to φ and at all external points is zero. Since the values of φ and its derivatives at all interior points are uniquely determined by the value of φ_S alone on S (Dirichlet Problem), or by $(\partial\varphi/\partial n)_S$ alone (Neumann Problem), the functions φ_S and $(\partial\varphi/\partial n)_S$ cannot be chosen independently if φ and u are to be identical at interior points. The function defined by (2) satisfies (1) and is regular within the domain V whatever the choice of φ_S and $(\partial\varphi/\partial n)_S$, but the values assumed by u and $\partial u/\partial n$ on S will in general differ from those assigned to φ_S and $(\partial\varphi/\partial n)_S$.

A Kirchhoff diffraction problem is formulated usually somewhat as follows. A primary wave is incident upon an opaque screen in which there is an opening S_1. The scalar potential of the wave, or a rectangular component of a field vector, is represented by φ satisfying (1). It is now assumed that on the dark side of the screen φ_S and $(\partial\varphi/\partial n)_S$ are zero except over the opening S_1, where φ_S and $(\partial\varphi/\partial n)_S$ have their undisturbed values. The diffracted wave is then calculated from (2), the integral extending over S_1 alone.

Such a procedure cannot possibly give an exact result. For in the first place the assumption of zero values for the light function and its derivative over the screen S_2 implies a discontinuity about the contour C_1 bounding the opening, while Green's theorem, upon which (2) is based, is valid only for functions which are continuous over a complete bounding surface. This difficulty cannot be obviated by the common expedient of replacing the contour of discontinuity by a thin region of rapid but continuous transition. The vanishing of φ_S and $(\partial\varphi/\partial n)_S$ on any finite part of S would then entail a zero value everywhere. The integral represents in fact a scalar wave function which approximates the true intensity to a degree which must be determined by other means. If the ratio of wave-length to size of S_1 is small, the radiation is thrown largely forward and the Kirchhoff function may differ by a negligible amount from the assumed zero value over the screen. If, on the other hand, the wave-length is large, the wave function calculated on the basis of undisturbed values over S_1 will be found to have values over S_2 which are by no means small. It is sometimes suggested[1] that this distribution over S_2 be applied once more to the Kirchhoff formula (2)—a method of successive approximations. Quite apart from questions of convergence, the difficulties of evaluating the resulting surface integrals in most cases make such a procedure of little practical value.

There are other difficulties. An electromagnetic field at a point on the closed surface S is characterized by a set of scalar functions which represent the rectangular components of the electric and magnetic vectors. Each one of these scalar functions satisfies (1) and its value at an interior point x', y', z' is therefore expressed by (2) in terms of its values over the boundary S. But these components at an interior point must not only satisfy the wave equation, *they must also be solutions of the Maxwell field equations.* The real problem, therefore, is not the integration of a scalar wave equation, nor even a vector wave equation, but of a simultaneous system of first-order vector equations relating the vectors E and H. There is nothing new in this remark. The integration of the field equations in closed form has been discussed by Love,[2] Larmor,[3] v. Ignatowsky,[4] Tonolo,[5] Macdonald,[6] Tedone,[7] and most completely by Kottler,[8] but their results appear to have been commonly disregarded in treatises on physical optics. Recently, however, the subject has been reviewed by Schelkunoff[9] in connection with equivalence theorems.

3. A Vector Analog of Green's Theorem

The integration of the field equations can be achieved most directly and rigorously by a method which is wholly analogous to the treatment of the scalar wave equation. Let V be a closed region of space bounded by a regular surface S, and let $\mathbf{P}$ and $\mathbf{Q}$ be two functions of position which together with their first and second derivatives are continuous throughout V and on the surface S. The divergence theorem is applied to the vector $\mathbf{P}\times\nabla\times\mathbf{Q}$, giving

$$\int_V \nabla\cdot(\mathbf{P}\times\nabla\times\mathbf{Q})dv = \int_S (\mathbf{P}\times\nabla\times\mathbf{Q})\cdot\mathbf{n}da, \quad (3)$$

where $\mathbf{n}$ is a unit normal vector directed outward from S. Upon expanding the integrand of the volume integral a vector analog of Green's first identity is obtained,

[1] M. Born, *Optik*, p. 152.

[2] A. E. H. Love, Phil. Trans. **A197**, 1 (1901).
[3] J. Larmor, Lond. Math. Soc. Proc. **1**, 1 (1903).
[4] W. v. Ignatowsky, Ann. d. Physik **23**, 875 (1907); **25**, 99 (1908).
[5] A. Tonolo, Annali di Matematica **3**, 17, 29 (1910).
[6] H. M. Macdonald, Proc. Lond. Math. Soc. **10**, 91 (1911); and Phil. Trans. **A212**, 295 (1912).
[7] O. Tedone, Linc. Rendi (5), **1**, 286 (1917).
[8] F. Kottler, Ann. d. Physik **71**, 457 (1923).
[9] S. A. Schelkunoff, Bell Sys. Tech. J. **15**, 92 (1936).

$$\int_V (\nabla\times\mathbf{P}\cdot\nabla\times\mathbf{Q}-\mathbf{P}\cdot\nabla\times\nabla\times\mathbf{Q})dv = \int_S (\mathbf{P}\times\nabla\times\mathbf{Q})\cdot\mathbf{n}da. \quad (4)$$

The vector analog of Green's second identity ("Green's theorem") is obtained by reversing the roles of $\mathbf{P}$ and $\mathbf{Q}$ in (4) and subtracting one expression from the other.

$$\int_V (\mathbf{Q}\cdot\nabla\times\nabla\times\mathbf{P}-\mathbf{P}\cdot\nabla\times\nabla\times\mathbf{Q})dv = \int_S (\mathbf{P}\times\nabla\times\mathbf{Q}-\mathbf{Q}\times\nabla\times\mathbf{P})\cdot\mathbf{n}da. \quad (5)$$

The first identity (4) is the logical basis for uniqueness proofs in connection with vector fields. It will be noted that if one place $\mathbf{P}=\mathbf{Q}=\mathbf{E}$, (4) proves to be identical with Poynting's theorem.

4. Application to the Field Equations

The field equations in a rationalized m.k.s. system of units are

$$\text{(I)}\quad \nabla\times\mathbf{E}-i\omega\mu\mathbf{H}=-\mathbf{J}^*, \qquad \text{(III)}\quad \nabla\cdot\mathbf{H}=\rho^*/\mu,$$

$$\text{(II)}\quad \nabla\times\mathbf{H}+i\omega\epsilon\mathbf{E}=\mathbf{J}, \qquad \text{(IV)}\quad \nabla\cdot\mathbf{E}=\rho/\epsilon.$$

It is assumed here that the fields are harmonic and that all quantities contain the time in the form of a factor $\exp(-i\omega t)$.

The field intensity $\mathbf{E}$ is measured in volts per meter, $\mathbf{H}$ in ampere-turns per meter, current density $\mathbf{J}$ in amperes per square meter. In free space $\mu_0=4\pi\times10^{-7}$ henry per meter, $\epsilon_0=(1/36\pi)\times10^{-9}$ farad per meter. The medium is assumed to be homogeneous and isotropic, and of zero conductivity. The quantities $\mathbf{J}^*$ and ρ^* are fictitious densities of magnetic current and magnetic charge. In normal fields they are zero, but the arbitrary (and physically impossible) assumption of discontinuities in the tangential components of $\mathbf{E}$ about the contour of an opening can only be accounted for by some such assumption. Currents and charges are related by the equations of continuity,

$$\text{(V)}\quad \nabla\cdot\mathbf{J}-i\omega\rho=0, \qquad \nabla\cdot\mathbf{J}^*-i\omega\rho^*=0.$$

The vectors $\mathbf{E}$ and $\mathbf{H}$ satisfy

$$\nabla\times\nabla\times\mathbf{E}-k^2\mathbf{E}=i\omega\mu\mathbf{J}-\nabla\times\mathbf{J}^*, \quad (6)$$

$$\nabla\times\nabla\times\mathbf{H}-k^2\mathbf{H}=i\omega\epsilon\mathbf{J}^*+\nabla\times\mathbf{J}, \quad (7)$$

where $k^2=\omega^2\epsilon\mu$. Solutions of (6) and (7) are to be found which are finite and single-valued at all interior points of V and at all points on S.

In (5) let $\mathbf{P}=\mathbf{E}$ and $\mathbf{Q}=\psi\mathbf{a}$, where $\mathbf{a}$ is a unit vector in an arbitrary direction and $\psi=e^{ikr}/r$. The Green's function $\mathbf{Q}$ is essentially the vector potential of a unit current element. Distance is measured from the element at x, y, z to the point of observation at x', y', z'.

$$r=[(x'-x)^2+(y'-y)^2+(z'-z)^2]^{\frac{1}{2}}. \quad (8)$$

The following identities are readily verified.

$$\nabla\times\mathbf{Q}=\nabla\psi\times\mathbf{a}, \qquad \nabla\times\nabla\times\mathbf{Q}=\mathbf{a}k^2\psi+\nabla(\mathbf{a}\cdot\nabla\psi),$$
$$\nabla\times\nabla\times\mathbf{P}=k^2\mathbf{E}+i\omega\mu\mathbf{J}-\nabla\times\mathbf{J}^*.$$

By application of the divergence theorem and further transformation, it is easily shown that $\mathbf{a}\cdot$ is a factor common to all the terms in (5), and since the direction of $\mathbf{a}$ is arbitrary, it follows that

$$\int_V\left[i\omega\mu\mathbf{J}\psi-\nabla\times\mathbf{J}^*\psi+\frac{1}{\epsilon}\rho\nabla\psi\right]dv = \int_S [i\omega\mu(\mathbf{n}\times\mathbf{H})\psi+(\mathbf{n}\times\mathbf{E})\times\nabla\psi + (\mathbf{n}\cdot\mathbf{E})\nabla\psi-\mathbf{n}\times\mathbf{J}^*\psi]da. \quad (9)$$

An application of the identity

$$\int_V \nabla\times\mathbf{J}^*\psi dv=\int_S \mathbf{n}\times\mathbf{J}^*\psi da+\int_V \mathbf{J}^*\times\nabla\psi dv \quad (10)$$

reduces this to

$$\int_V\left[i\omega\mu\mathbf{J}\psi-\mathbf{J}^*\times\nabla\psi+\frac{1}{\epsilon}\rho\nabla\psi\right]dv = \int_S [i\omega\mu(\mathbf{n}\times\mathbf{H})+(\mathbf{n}\times\mathbf{E})\times\nabla\psi + (\mathbf{n}\cdot\mathbf{E})\nabla\psi]da. \quad (11)$$

Because of the singularity of the function ψ at $r=0$, the identity (11) holds only when this point is excluded from V. The procedure is familiar to everyone acquainted with potential

theory. The point x', y', z' is taken as the center of a small sphere of radius r_1. The normal over the sphere is directed out of V, and consequently radially *towards* the center.

$$\nabla\psi=(1/r-ik)(e^{ikr}/r)\mathbf{r}_0, \tag{12}$$

and on the sphere $\mathbf{n}=\mathbf{r}_0$. The area of the sphere vanishes with the radius as $4\pi r_1^2$, and since

$$(\mathbf{n}\times\mathbf{E})\times\mathbf{n}+(\mathbf{n}\cdot\mathbf{E})\mathbf{n}=\mathbf{E}, \tag{13}$$

the contribution of the spherical surface to the right-hand side of (11) reduces to $4\pi\mathbf{E}(x', y', z')$. The value $\mathbf{E}$ at any internal point of V is therefore

$$\begin{aligned}\mathbf{E}(x', y', z') &=\frac{1}{4\pi}\int_V\left[i\omega\mu\mathbf{J}\psi-\mathbf{J}^*\times\nabla\psi+\frac{1}{\epsilon}\rho\nabla\psi\right]dv\\ &\quad-\frac{1}{4\pi}\int_S[i\omega\mu(\mathbf{n}\times\mathbf{H})\psi\\ &\quad+(\mathbf{n}\times\mathbf{E})\times\nabla\psi+(\mathbf{n}\cdot\mathbf{E})\nabla\psi]da.\end{aligned} \tag{14}$$

An appropriate interchange of vectors gives

$$\begin{aligned}\mathbf{H}(x', y', z') &=\frac{1}{4\pi}\int_V\left[i\omega\epsilon\mathbf{J}^*\psi+\mathbf{J}\times\nabla\psi+\frac{1}{\mu}\rho^*\nabla\psi\right]dv\\ &\quad+\frac{1}{4\pi}\int_S[i\omega\epsilon(\mathbf{n}\times\mathbf{E})\psi-(\mathbf{n}\times\mathbf{H})\times\nabla\psi\\ &\quad-(\mathbf{n}\cdot\mathbf{H})\nabla\psi]da.\end{aligned} \tag{15}$$

It will be shown below under more general circumstances that (14) and (15) satisfy the field equations at all points of V and S.

These expressions are essentially equivalent to those obtained by v. Ignatowsky for a closed surface. If all sources can be enclosed within a sphere of finite radius, the field is regular at infinity and either side of S may be chosen as its "interior," or S may be closed at infinity. It may be remarked that (14) and (15) are convenient expressions for the calculation of a field directly from a given distribution of current without the intervention of vector and scalar potentials or of a Hertzian vector. The surface S recedes to infinity and the fictitious magnetic sources are placed equal to zero. Then

$$\begin{aligned}\mathbf{E}(x', y', z') &=\frac{1}{4\pi}\int_V\left[i\omega\mu\mathbf{J}\psi+\frac{1}{\epsilon}\rho\nabla\psi\right]dv,\\ \mathbf{H}(x', y', z') &=\frac{1}{4\pi}\int_V\mathbf{J}\times\nabla\psi dv.\end{aligned} \tag{16}$$

Since the current distribution is given, the charge density can be determined from the equation of continuity.

It is known from the uniqueness theorem of electromagnetic theory that a field within a bounded domain is completely determined by the specification of the tangential components of $\mathbf{E}$ and $\mathbf{H}$ on the surface. It follows in (14) and (15) that when $\mathbf{n}\times\mathbf{E}$ and $\mathbf{n}\times\mathbf{H}$ have been fixed, the choice of $\mathbf{n}\cdot\mathbf{E}$ and $\mathbf{n}\cdot\mathbf{H}$ is no longer arbitrary. The selection must be consistent with the conditions on a field satisfying Maxwell's equations. The same limitations on the choice of φ_S and $(\partial\varphi/\partial n)_S$ were pointed out in §2. The dependence of the normal component of $\mathbf{E}$ upon the tangential component of $\mathbf{H}$ is equivalent to that of ρ upon $\mathbf{J}$.

Let us suppose for the moment that the charge and current distributions in (14) are confined to a thin layer at the surface S. As the depth of the layer diminishes the densities may be increased so that in the limit the volume densities are replaced in the usual way by surface densities. If the region V contains no charge or current within its interior or on its boundary S, the field at an interior point is

$$\begin{aligned}\mathbf{E}(x', y', z') &=\frac{1}{4\pi}\int_S[i\omega\mu(\mathbf{n}\times\mathbf{H})\psi\\ &\quad+(\mathbf{n}\times\mathbf{E})\times\nabla\psi+(\mathbf{n}\cdot\mathbf{E})\nabla\psi]da.\end{aligned} \tag{17}$$

It is now clear that this is exactly the field that would be produced by a distribution of electric current over S with surface density $\mathbf{K}$, a distribution of magnetic current of density $\mathbf{K}^*$, and a surface electric charge of density η, where

$$\mathbf{K}=-\mathbf{n}\times\mathbf{H}, \qquad \mathbf{K}^*=\mathbf{n}\times\mathbf{E}, \qquad \eta=-\epsilon\mathbf{n}\cdot\mathbf{E}. \tag{18}$$

The values of $\mathbf{E}$ and $\mathbf{H}$ in (18) are those just *inside* the surface S. The function $\mathbf{E}(x', y', z')$ defined by (17) is discontinuous across S. It is

well known[10] that the integral

$$\mathbf{E}_3(x', y', z') = \frac{1}{4\pi\epsilon}\int_S \eta\nabla\psi da \qquad (19)$$

suffers a discontinuity on transition through S given by $\mathbf{n}\cdot\Delta\mathbf{E}_3 = \eta/\epsilon$, where $\Delta\mathbf{E}_3$ is the difference of the values outside and inside. The third term of (17) does not affect the transition of the tangential component but reduces the normal component to zero. Likewise the discontinuity of

$$\mathbf{E}_2(x', y', z') = \frac{1}{4\pi}\int_S \mathbf{K}^*\times\nabla\psi da \qquad (20)$$

is specified by $\mathbf{n}\times\mathbf{E}_2 = \mathbf{K}^*$, so that the second term in (17) reduces the tangential component of $\mathbf{E}$ to zero without affecting the normal component. The first term in (17) is continuous across S, but has discontinuous derivatives.

$$\nabla'\times\mathbf{E}(x', y', z') = -\frac{i\omega\mu}{4\pi}\int_S (\mathbf{n}\times\mathbf{H})\times\nabla\psi da. \qquad (21)$$

The vector $\mathbf{E}$ and the tangential component of its curl is zero on the positive side of S; it is therefore zero at all external points. The same analysis applies to $\mathbf{H}$.

5. Extension to Discontinuous Surface Distributions

The results of the preceding section hold only if the vectors $\mathbf{E}$ and $\mathbf{H}$ are continuous and have continuous first derivatives at all points of S. They cannot, therefore, be applied directly to the problem of diffraction at a slit. To obtain the required extension of (17) to such cases, consider the closed surface S (surfaces closed at infinity are included) to be divided into two zones S_1 and S_2 by a closed contour C lying on S, as in Fig. 1. The vectors $\mathbf{E}$ and $\mathbf{H}$ and their first derivatives are continuous over S_1 and satisfy the field equations. The same is true over S_2. However the components of $\mathbf{E}$ and $\mathbf{H}$ which are tangential to the surface are now subject to a discontinuous change in passing across C from one zone to the other. The occurrence of such discontinuities can be reconciled with the field equations only by the further assumption of a line distribution of charges or currents about the contour C. This line distribution of sources contributes to the field and only when it is taken into account do the resultant expressions for $\mathbf{E}$ and $\mathbf{H}$ satisfy Maxwell's equations. The calculation of a diffraction pattern from an integral extended over a portion of a surface only, as has been the customary practice, must necessarily lead to erroneous results.

A method of determining a contour distribution consistent with the requirements of the problem was proposed by Kottler.[8] It was shown in §4 that the field at an interior point in (17) is identical with that produced by the surface currents and charges specified in (18). A discontinuity in the tangential components of $\mathbf{E}$ and $\mathbf{H}$ in passing on the surface from zone S_1 to zone S_2 implies therefore an abrupt change in the surface current density. The termination of a line of current, in turn, can be accounted for on the basis of the equation of continuity by an accumulation of charge on the contour. Let $d\mathbf{s}$ be an element of length along the contour in the positive direction as determined by the positive normal $\mathbf{n}$, Fig. 1. Let $\mathbf{n}_1$ be a unit vector lying in the surface, normal to both $\mathbf{n}$ and the contour element $d\mathbf{s}$, and directed into zone (1). The line densities of electric and magnetic charge will be designated by σ and σ^*. Then Eqs. (V), when applied to surface currents, become

$$\mathbf{n}_1\cdot(\mathbf{K}_1-\mathbf{K}_2) = i\omega\sigma, \quad \mathbf{n}_1\cdot(\mathbf{K}_1^*-\mathbf{K}_2^*) = i\omega\sigma^*, \qquad (22)$$

and hence by (18),

$$i\omega\sigma = \mathbf{n}_1\cdot(\mathbf{n}\times\mathbf{H}_2-\mathbf{n}\times\mathbf{H}_1) = (\mathbf{H}_2-\mathbf{H}_1)\cdot(\mathbf{n}_1\times\mathbf{n}), \qquad (23)$$

$$i\omega\sigma^* = \mathbf{n}_1\cdot(\mathbf{n}\times\mathbf{E}_1-\mathbf{n}\times\mathbf{E}_2) = -(\mathbf{E}_2-\mathbf{E}_1)\cdot(\mathbf{n}_1\times\mathbf{n}).$$

The vector $\mathbf{n}_1\times\mathbf{n}$ is in the direction of $d\mathbf{s}$. If S_2 represents an opaque screen over which $\mathbf{E}_2 = \mathbf{H}_2 = 0$, the field at any point on the shadow side is

$$\mathbf{E}(x', y', z') = -\frac{1}{i\omega\epsilon}\frac{1}{4\pi}\oint_C \nabla\psi\mathbf{H}_1\cdot d\mathbf{s} - \frac{1}{4\pi}\int_{S_1}[i\omega\mu(\mathbf{n}\times\mathbf{H}_1)\psi+(\mathbf{n}\times\mathbf{E}_1)\times\nabla\psi + (\mathbf{n}\cdot\mathbf{E}_1)\nabla\psi]da, \qquad (24)$$

[10] H. B. Phillips, *Vector Analysis* (Wiley and Sons, 1933), p. 206. In greater detail, H.Poincaré, *Théorie Mathématique de la Lumière*, II, Ch. VII.

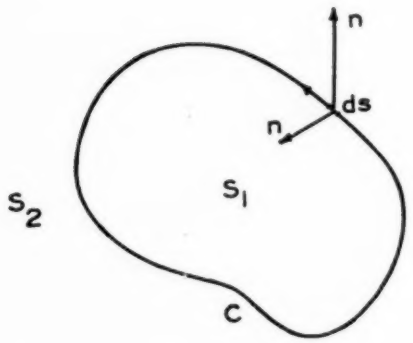


FIG. 1. The contour on which the fields and their derivatives are discontinuous. The lower **n** lying within the zone S_1 should be $\mathbf{n}_1$.

which can be shown to be identical with Kottler's result,

$$4\pi\mathbf{E}(x', y', z') = -\frac{1}{i\omega\epsilon}\oint_C \nabla\psi\mathbf{H}_1\cdot d\mathbf{s} + \oint_C \psi\mathbf{E}_1\times d\mathbf{s} - \int_{S_1}\left(\mathbf{E}_1\frac{\partial\psi}{\partial n} - \psi\frac{\partial\mathbf{E}_1}{\partial n}\right)da. \quad (25)$$

For the magnetic field one obtains

$$\begin{aligned}4\pi\mathbf{H}(x', y', z') &= \frac{1}{i\omega\mu}\oint_C \nabla\psi\mathbf{E}_1\cdot d\mathbf{s} + \int_{S_1}[i\omega\epsilon(\mathbf{n}\times\mathbf{E}_1)\psi \\ &\quad -(\mathbf{n}\times\mathbf{H}_1)\times\nabla\psi - (\mathbf{n}\cdot\mathbf{H}_1)\nabla\psi]da \\ &= \frac{1}{i\omega\mu}\oint_C \nabla\psi\mathbf{E}_1\cdot d\mathbf{s} + \oint_C \psi\mathbf{H}_1\times d\mathbf{s} \\ &\quad -\int_{S_1}\left(\mathbf{H}_1\frac{\partial\psi}{\partial n} - \psi\frac{\partial\mathbf{H}_1}{\partial n}\right)da. \quad (26)\end{aligned}$$

It remains to be shown that the fields expressed by these integrals are in fact divergenceless and satisfy (I) and (II). Consider first the divergence of (24) at a point x', y', z'.

$$\begin{aligned}\nabla'\cdot E(x', y', z') &= \frac{1}{i\omega\epsilon}\frac{1}{4\pi}\oint_C \nabla^2\psi\mathbf{H}_1\cdot d\mathbf{s} + \frac{1}{4\pi}\int_S [i\omega\mu(\mathbf{n}\times\mathbf{H}_1)\cdot\nabla\psi \\ &\quad +(\mathbf{n}\cdot\mathbf{E}_1)\nabla^2\psi]da \\ &= -\frac{k^2}{i\omega\epsilon}\frac{1}{4\pi}\oint_C \psi\mathbf{H}_1\cdot d\mathbf{s} - \frac{k^2}{4\pi}\int_{S_1}(\mathbf{n}\cdot\mathbf{E}_1)\psi da \\ &\quad +\frac{i\omega\mu}{4\pi}\int_{S_1}(\mathbf{n}\times\mathbf{H}_1)\cdot\nabla\psi da, \quad (27)\end{aligned}$$

taking into account the relation $\nabla' = -\nabla$ when applied to ψ or its derivatives. Now

$$\int_S (\mathbf{n}\times\mathbf{H})\cdot\nabla\psi da = \int_S \psi\nabla\times\mathbf{H}\cdot\mathbf{n}da - \oint_C \psi\mathbf{H}\cdot d\mathbf{s}. \quad (28)$$

The line integral resulting from this transformation is zero when S is closed, but otherwise just cancels the contour integral in (27). Inversely, only the presence of the contour integral leads to a zero divergence and transverse waves at great distances from the opening S_1. From (II) and the relation $k^2/i\omega\epsilon = -i\omega\mu$ follows immediately the result

$$\nabla'\cdot\mathbf{E}(x', y', z') = 0. \quad (29)$$

An identical proof holds for $\mathbf{H}(x', y', z')$.

Finally it will be shown that (24) and (26) satisfy (I) and (II).

$$\nabla'\times\mathbf{H}(x', y', z') = -\frac{1}{4\pi}\int_{S_1}[(\mathbf{n}\times\mathbf{H}_1)\cdot\nabla\nabla\psi + k^2\psi(\mathbf{n}\times\mathbf{H}_1) - i\omega\epsilon(\mathbf{n}\times\mathbf{E}_1)\times\nabla\psi]da, \quad (30)$$

since the curl of the gradient is identically zero. Furthermore

$$\begin{aligned}\int_{S_1}(\mathbf{n}\times\mathbf{H}_1)\cdot\nabla\nabla\psi da &= -\int_{S_1}(\mathbf{n}\cdot\nabla_{\nabla\psi}\times\mathbf{H}_1)\nabla\psi da \\ &= \int_{S_1}(\mathbf{n}\cdot\nabla\times\mathbf{H}_1)\nabla\psi da - \int_{S_1}\mathbf{n}\times\nabla\cdot(\mathbf{H}_1\nabla\psi)da \\ &= -\oint_C \nabla\psi\mathbf{H}_1\cdot d\mathbf{s} - i\omega\epsilon\int_{S_1}(\mathbf{n}\cdot\mathbf{E}_1)\nabla\psi da, \quad (31)\end{aligned}$$

where the operator $\nabla_{\nabla\psi}$ acts on $\nabla\psi$ only. The last integral takes account of the fact that the field equations are by hypothesis satisfied on S_1. Then

$$\begin{aligned}\nabla'\times\mathbf{H}(x', y', z') &= \frac{1}{4\pi}\oint_C \nabla\psi\mathbf{H}_1\cdot d\mathbf{s} + \frac{i\omega\epsilon}{4\pi}\int_{S_1}[i\omega\mu(\mathbf{n}\times\mathbf{H}_1)\psi \\ &\quad +(\mathbf{n}\times\mathbf{E}_1)\times\nabla\psi + (\mathbf{n}\cdot\mathbf{E}_1)\nabla\psi]da \\ &= -i\omega\epsilon\mathbf{E}(x', y', z'). \quad (32)\end{aligned}$$

The validity of (I) is established in the same manner.

6. Application to a Rectangular Slit

In Fig. 2 a rectangular slit is shown in an infinite, perfectly conducting screen coinciding with the xy plane. Plane waves coming from the left are incident on this slit at various angles. The normal to the slit surface is drawn to the left and the intensity of the field is to be calculated at any point whose coordinates are R', θ', φ' lying to the right of the screen. Then at sufficiently large distances

$$r \backsimeq R' - x \cos \varphi' \sin \theta' - y \sin \varphi' \sin \theta'. \quad (33)$$

Consider first the case of a normal incidence, with the electric vector polarized along the x axis and of unit intensity. The field at great distances from the slit obtained by evaluating (25) is

$$\begin{aligned} 4\pi E_{\theta'}' &= -ik(1+\cos\theta')\cos\varphi' A, \\ 4\pi E_{\varphi'}' &= ik(1+\cos\theta')\sin\varphi' A, \\ E_{R'}' &= 0, \end{aligned} \quad (34)$$

where

$$A = 4\frac{\sin(\frac{1}{2}ka\cos\varphi'\sin\theta')}{k\cos\varphi'\sin\theta'} \times \frac{\sin(\frac{1}{2}kb\sin\varphi'\sin\theta')}{k\sin\varphi'\sin\theta'}\frac{e^{ikR'}}{R'}. \quad (35)$$

Now if this solution is extended analytically towards the screen, it is evident that it does not vanish on the plane $z=0$, as required by the boundary conditions. In fact nothing has been stated as to the location of the surface S_2 which closes the slit surface S_1, and one is free to choose it in the way which is least liable to violate the actual conditions. Thus, in the present case S_1 may be closed by a surface S_2 which lies just behind S_1 and over which it is assumed that the field vectors vanish. This is equivalent to saying that the field functions (34) can be continued analytically into the region for which $\theta' > \pi/2$. The effect of the screen or baffle is now taken into account by assuming it to act as a perfect reflector. The phase relations after reflection are determined by the condition that the resultant

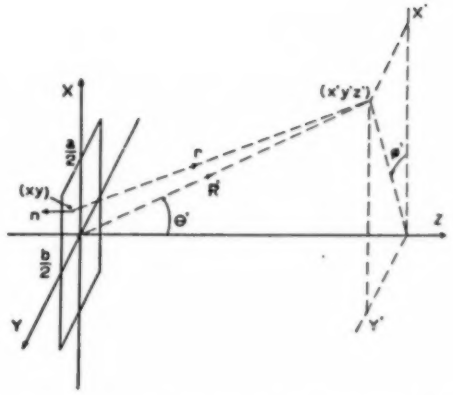


Fig. 2. Coordinate system for calculating the diffraction by a rectangular slit.

tangential component of **E** must be zero at $\theta' = \pi/2$, while the magnitude of the initial normal component is doubled. In the present instance these conditions are expressed by

$$\begin{aligned} E_{\theta'} &= E_{\theta'}'(\theta') + E_{\theta'}'(\pi-\theta'), \\ E_{\varphi'} &= E_{\varphi'}'(\theta') - E_{\varphi'}'(\pi-\theta'), \end{aligned} \quad (36)$$

where $E_{\theta'}$ and $E_{\varphi'}$ are normal and tangential components of the *resultant* field. Applied to (34) this gives

$$\begin{aligned} 4\pi E_{\theta'} &= -2ik\cos\varphi' A, \\ 4\pi E_{\varphi'} &= 2ik\cos\theta'\sin\varphi' A. \end{aligned} \quad (37)$$

The energy flow or intensity of the diffracted wave is proportional to the sum of the squares of these quantities. The solid curves of Fig. 3 are plots of the intensity in the vertical plane $\varphi' = 0$ for several ratios of slit breadth to wavelength. Similar curves are drawn in Fig. 6 for the horizontal plane $\varphi' = \pi/2$.

In the more general case the direction of propagation of the incident wave makes an arbitrary angle α with the z axis. If the polarization is parallel to the x axis, the components of the incident wave are

$$E_x = e^{ik(y\sin\alpha + z\cos\alpha)}, \quad H_y = (\epsilon/\mu)^{\frac{1}{2}}\cos\alpha E_x, \quad H_z = -(\epsilon/\mu)^{\frac{1}{2}}\sin\alpha E_x. \quad (38)$$

The diffracted field is found from (25) to be

$$\begin{aligned} 4\pi E_{\theta'}' &= -ik\cos\varphi'(1+\cos\theta'\cos\alpha)A, \\ E_{R'}' &= 0 \\ 4\pi E_{\varphi'}' &= ik\sin\varphi'(\cos\theta'+\cos\alpha)A, \end{aligned} \quad (39)$$

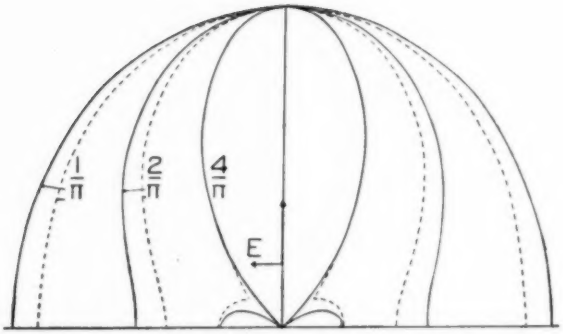


FIG. 3.

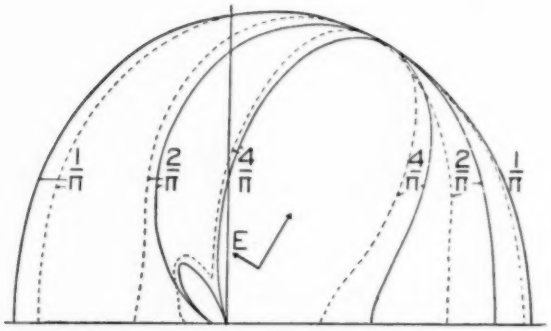


FIG. 4.

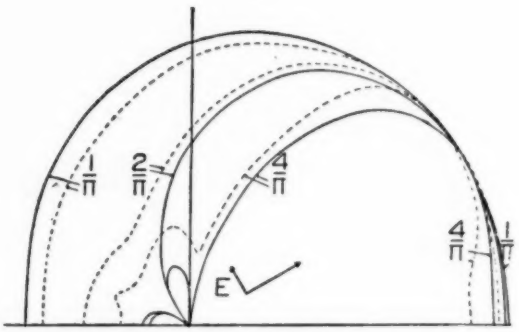


FIG. 5.

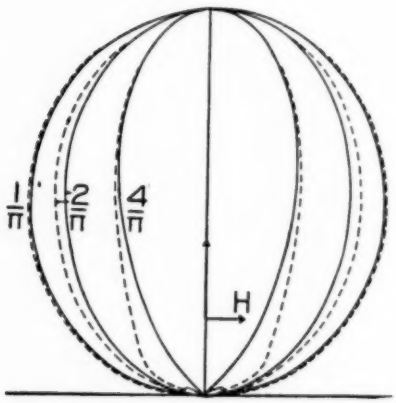


FIG. 6.

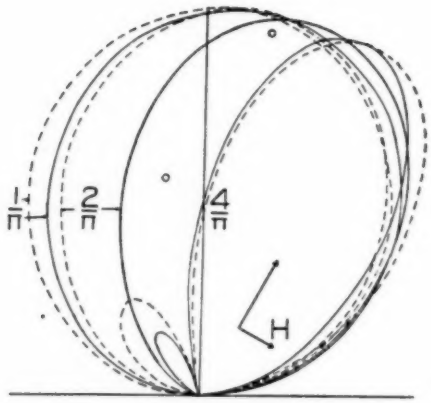


FIG. 7.

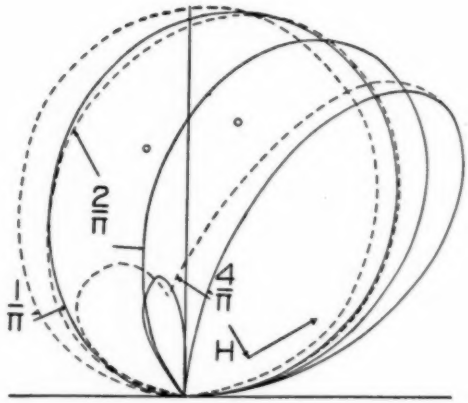


FIG. 8.

FIGS. 3–8. Polar diagrams of the distribution of field intensity, showing the diffraction of a plane wave by a slit in the plane of incidence. The long arrow indicates the direction of incidence and the short arrow indicates the polarization of the field at the slit. The numerals on the curves indicate the width of the slit in wave-lengths (a/λ or b/λ).

and the resultant field to the right of the plane screen obtained by adding the reflected wave is identical with (37), but with α now entering into A in the form

$$A=4\frac{\sin\left(\frac{1}{2}ka\cos\varphi'\sin\theta'\right)}{k\cos\varphi'\sin\theta'}\times\frac{\sin\frac{1}{2}kb(\sin\varphi'\sin\theta'-\sin\alpha)}{k(\sin\varphi'\sin\theta'-\sin\alpha)}\frac{e^{ikR'}}{R'}. \quad (40)$$

In Figs. 7 and 8 the solid curves represent plots of intensity in the horizontal yz plane for $\alpha=30°$ and $\alpha=60°$. The distribution in the vertical plane is identical with Fig. 3.

In the alternate case the magnetic vector is

polarized vertically along the x axis and the incident wave is defined by

$$H_x = e^{ik(y\sin\alpha + x\cos\alpha)}, \quad E_y = -(\mu/\epsilon)^{\frac{1}{2}}\cos\alpha H_x, \qquad E_z = (\mu/\epsilon)^{\frac{1}{2}}\sin\alpha H_x. \tag{41}$$

The diffracted field calculated from (25) is then

$$4\pi E_{\theta'}' = i\omega\mu\sin\varphi'(\cos\theta' + \cos\alpha)A, \qquad 4\pi E_{\varphi'}' = i\omega\mu\cos\varphi'(1+\cos\theta\cos\alpha)A, \tag{42}$$

and the resultant field after reflection is

$$4\pi E_{\theta'} = 2i\omega\mu\sin\varphi'\cos\alpha A, \qquad 4\pi E_{\varphi'} = 2i\omega\mu\cos\varphi'\cos\theta'\cos\alpha A, \tag{43}$$

where A is defined by (40). The solid curves of Figs. 4 and 5 represent plots of intensity in the horizontal plane. The distribution in the vertical plane is identical with Fig. 6.

It is exceedingly interesting to compare these results with some calculations made recently by Morse and Rubenstein,[11] who have carried through the two-dimensional problem of diffraction of an electromagnetic plane wave by an infinite slit. The two methods should lead to approximately the same distribution in an equatorial plane. The intensity plots obtained from this rigorous solution are shown as dotted curves. The correspondence on the whole is remarkably good. It will be noted that the most marked deviation occurs in the immediate neighborhood of the screen and probably arises from the errors which are fundamental to the present method: the assumption of unperturbed distributions over the slit and the manner in which the reflection problem has been handled. The authors have discarded two points in Figs. 7 and 8 which fall on the curves published by Morse and Rubenstein[11] and which lead to lobes for which it is difficult to account. This neglect may or may not be justified. Professor Morse has kindly put the original data at the disposition of the authors but insufficient calculations were made to settle the matter one way or the other. It is hard to reconcile the anomalous occurrence of these lobes with the close correspondence of all other points. Whatever the answer to this question, the results give strong support to the belief that Eqs. (25) and (26) can be applied to the calculation of diffracted radiation with assurance of reasonable accuracy. In the case of radiation from hollow tubes and horns, an extension to take account of the internal reflected wave is no doubt possible based on the methods employed in acoustics under similar circumstances.

[11] P. M. Morse and Pearl J. Rubenstein, Phys. Rev. **54**, 895 (1938).

Some Physical Properties of Compressed Gases

VI. The Fugacity of Carbon Dioxide

W. Edwards Deming and Lola S. Deming
Bureau of Chemistry and Soils, U. S. Department of Agriculture, Washington
(Received May 5, 1939)

The Michels' compressibility data for carbon dioxide are smoothed and interpolated by graphical methods, and the fugacity is computed at certain integral pressures. The temperature range is from 0° to 150°C, and the pressure range from 0 to the neighborhood of 3000 atmos. Along all isotherms, the fugacity at first drops below the pressure, the ratio f/p decreasing with p—a phenomenon well known and explained by the attractive forces of the molecules. The minimum in f/p is reached at or near 600 atmos. for all isotherms above critical. With further increase in pressure, the ratio f/p increases, so that the fugacity eventually equals the pressure and thereafter exceeds it. At 50°, $f=p$ at 2435 atmos.; at 150°, $f=p$ at 1675 atmos.

1. The Data

THE Michels and their collaborators have published compressibility data and various thermodynamic properties of carbon dioxide.[1–3] For facilitating calculations of equilibria involving this gas, it has been thought advisable to add fugacities to the list.

The compressibility data[1–3] are published in the form of pv against density and pressure, in isotherms and isochores, the temperature range being from 0° to 150°, and the density range 18.75 to 596 Amagat, i.e., from roughly 0.8 to 27 moles per liter. At 27 moles per liter and 150°C, the pressure is 3240 atmos.

2. A Word on Amagat Units

In order to determine the density of the compressed gas in g/cc, it would be necessary for the experimenter to know not only the volume of the high pressure pipet, but also how many grams of gas there are in it. The custom since the time of Amagat has been to express the densities otherwise, namely, in a unit now called after him, and designating the ratio of the volume that the compressed gas *would* occupy if expanded to 1 atmos. and 0°, to the volume that it actually occupies in the high pressure pipet at pressure p and temperature t°C. Thus, if the Amagat density of a gas at a certain pressure and temperature is stated to be 20, we know that this same gas would occupy 20 times as much space at 1 atmos. and 0°. Amagat volumes are simply reciprocals of Amagat densities; the Amagat volume of the gas whose density is 20 is therefore 1/20th; i.e., its volume is only 1/20th of what it would be if it were expanded to 1 atmos. and 0°. The Amagat system circumvents the need for weighing the gas; nevertheless, if the density of the gas at 1 atmos. and 0° is known in g/cc from other sources, then it is simple to convert Amagat densities into g/cc, and Amagat volumes into cc/g. Or, if the pv product at some temperature and pressure is known in Amagat units and also in (e.g.) mol. atmos./cc (as is the case at zero pressure and 0°, *vide infra*), then the factor for converting pv and v from the Amagat system into cc atmos./mol. and cc/mol., respectively, is evident.

In practice, the experimental gas is not expanded from the high pressure pipet down to 1 atmos. and 0°, but instead, for convenience, to a pressure only near 1 atmos. and at 25°, wherefore it becomes necessary to translate the measurements at 25° into the standard temperature and pressure (1 atmos., 0°). This translation can be achieved through the following considerations.

Suppose that the relation[4]

$$pv=A+B/v+C/v^2 \qquad (36)$$

[1] A. and C. Michels and H. Wouters, Proc. Roy. Soc. **A153**, 201–224 (1935).
[2] C. Michels, Thesis (Amsterdam, 1937).
[3] A. Michels, A. Bijl, and C. Michels, Proc. Roy. Soc. **A160**, 376–384 (1937).

[4] The equations, figures, and tables are numbered as continuations from those in the previous papers of this series. The references are shown in the next footnote.

with the possible addition of a term D/v^4 represents the trend of pv at 25° for low pressure measurements (perhaps to 40 or 50 atmos.) on a certain amount of the gas, v being in any desired units, such as cc. Also, let a similar series with constants A_0, B_0, C_0 represent the trend at 0°. A and A_0, it will be noticed, are the pv products at infinite attenuation ($p=\rho=0$), hence if v_1 denotes the volume of the gas at 0° and 1 atmos., then the ratio $1+\lambda$ of the pv product at 0° and infinite attenuation to the pv product at 0° and 1 atmos. will be

$$1+\lambda=1/(1+B_0/A_0v_1+C_0/A_0v_1^2), \quad (37)$$

where

$$v_1=A_0+B_0/v_1+C_0/v_1^2. \quad (38)$$

Evidently $1+\lambda$ is the value of pv in Amagat units for 0° and infinite attenuation; it is independent of the amount of gas used, and of the units in which v is measured.

Now since at infinite attenuation, the pv product for a given mass of gas is proportional to the absolute temperature, we may write

$$(pv)_s=(pv)_{25}\frac{T_0}{T_0+25}\times\frac{1}{(1+B/Av+C/Av^2)_{25°}}\frac{1}{1+\lambda} \quad (39)$$

for the relation between the pv product of a given mass of the gas at 25° and some low pressure usually near 1 atmos., and the pv product *of the same gas* under standard conditions (0° and 1 atmos., denoted by the subscript s). T_0 denotes the ice point, which, following the Michels, we assume in this work to be 273.15°C.

In practice the gas is expanded from the high pressure pipet at pressure p and temperature t into a calibrated low pressure pipet at 25° (usually near 1 atmos.), whereupon $(pv)_{25}$ is determined; thence by the relation just written comes also $(pv)_s$, which is *numerically* the volume that this same gas would occupy under standard conditions; the ratio of the volume occupied at the high pressure p and temperature t to that which it would occupy under standard conditions is the required Amagat volume v at the high pressure p and temperature t. The product of p and v gives the recorded pv in Amagat units corresponding to p and t.

It is worthy of note that though A_0, B_0, and C_0 all appear in the expression for $1+\lambda$ (Eq. 37), they do so only through the small correction term B_0/A_0 and the still smaller one C_0/A_0; hence $1+\lambda$ is determined with much greater relative precision than any of the parameters A_0, B_0, and C_0. In this work on carbon dioxide we found from the seven low pressure points at 0° cited in the first footnote,

$$1+\lambda=1.00679,$$

the root mean square variation being 0.00005 in the seven residuals, there being 7–3 degrees of freedom. The Michels used 1.006824, which is not significantly different.

Similar remarks hold for the precision of the denominator of Eq. (39), wherein A, B, and C for 25° also enter only as correction terms. It follows that with reasonably good compressibility data along the 0° and 25° isotherms, there should be negligible error in translating $(pv)_{25}$ into $(pv)_s$, and the units of pv should therefore be dependably Amagat.

3. Some Conversion Factors for Carbon Dioxide

Birge's value 22414.1±0.8 cc for the volume of a mol. of a perfect gas at 1 atmos. and 0° seems well established. The atomic weight of carbon is not so certain, but we adopt 12.010 which is more than close enough for any conversions that may be needed. With our value of $1+\lambda$ (*vide supra*) we can then say that

At $t=0°$, $p=0$,
$pv=RT_0=22414.1$ cc atmos./mol. (Birge)
$pv=1+\lambda=1.00679$ Amagat units
$pv=509.30$ cc atmos./g
At $t=0°$, $p=1$, $pv=1$ in Amagat units (definition)
$pv=22414.1/(1+\lambda)=22263.0$ cc atmos./mol.
$pv=22263/44.010=505.86$ cc atmos./g
$pv=22.4141/1000.03(1+\lambda)$
$=22.2623\ l$ atmos./mol.

By obvious divisions it follows that to convert (i) Amagat densities into mol./cc, divide them by 22263.0, or multiply them by 0.0000449177.

TABLE VI. *The fugacity of carbon dioxide at integral pressures. * denotes extrapolation; (s) saturation.*

p ATMOS.	ρ MOL./LITER	LOG f/p	f ATMOS.
	$t=0°$		
1	0.044919	$\bar{1}$.99684	0.99275
25	1.39	.9211	20.8
34.12(*s*)	2.19	.8882	26.4
	$t=5°$		
1	.044094	$\bar{1}$.99707	0.99328
25	1.34	.9266	21.1
38.30(*s*)	2.50	.8807	29.1
	$t=10°$		
1	.043300	$\bar{1}$.99724	0.99367
25	1.30	.9308	21.3
44.32(*s*)	3.05	.8675	32.7
	$t=15°$		
1	.042534	$\bar{1}$.99741	0.99405
25	1.26	.9352	21.5
50	3.60	.8580	36.1
50.50(*s*)	3.69	.8561	36.3
	$t=20°$		
1	.041796	$\bar{1}$.99754	0.99435
25	1.22	.9385	21.7
50	3.28	.8579	36.0
56.59(*s*)	4.43	.8454	39.6
	$t=25°$		
1	.041083	$\bar{1}$.99770	0.99472
25	1.19	.9425	21.9
50	3.05	.8770	37.7
63.47(*s*)	5.50	.8343	43.3
	$t=30°$		
1	.040394	$\bar{1}$.99792	0.99522
25	1.16	.9480	22.2
50	2.88	.8860	38.5
71.16(*s*)	7.75	.8155	46.5
	$t=31.04°$		
1	0.040263	$\bar{1}$.99784	0.99504
25	1.15	.946	22.1
50	2.86	.886	38.5
75	14.6	.810	48.4
100	17.4	.712	51.5
150	19.0	.573	56.1
200	20.1	.498	63.0
300	21.0	.414	77.8
400	22.2	.372	94.2
500	23.1	.348	111
600	23.7	.336	130
700	24.3	.335	151
800	24.7	.343	176
900	25.1	.357	205
1000	25.5	.379	239
*1500	27.1	.549	531
	$t=32°$		
1	.040126	$\bar{1}$.99788	0.99513
25	1.14	.947	22.1
50	2.82	.890	38.8
75	12.7	.816	49.1
100	17.1	.720	52.5
150	19.0	.578	56.8
200	20.1	.495	62.6
300	21.3	.416	78.2
400	22.2	.375	94.9
500	23.1	.353	113
600	23.6	.345	133
700	24.2	0.349	157
800	24.7	.361	184
900	25.2	.377	214
1000	25.5	.398	250
1500	27.0	.552	535
	$t=40°$		
1	.039085	$\bar{1}$.99816	0.99577
25	1.10	.954	22.5
50	2.62	.904	40.1
75	5.46	.837	51.5
100	14.4	.764	58.1
150	17.8	.639	65.3
200	19.1	.561	72.8
300	20.8	.469	88.3
400	21.8	.431	108
500	22.6	.409	128
600	23.3	.402	151
700	23.9	.405	178
800	24.5	.417	209
900	25.0	.433	244
1000	25.4	.454	284
1500	26.9	.593	588
	$t=50°$		
1	.037860	$\bar{1}$.99828	.99605
25	1.05	.957	22.7
50	2.43	.910	40.6
75	4.52	.858	54.1
100	9.13	.798	62.8
150	16.1	.675	71.1
200	17.9	.601	79.8
300	19.8	.516	98.4
400	21.1	.471	118
500	22.1	.452	142
600	22.8	.450	169
700	23.4	.454	199
800	23.8	.462	232
900	24.2	.478	271
1000	24.6	.497	314
1500	26.2	.641	657
*2000	27.6	.825	1337
	$t=75°$		
1	.035111	$\bar{1}$.99864	0.99688
25	0.950	.966	23.1
50	2.09	.931	42.7
75	3.53	.895	58.9
100	5.43	.858	72.1
150	10.9	.782	90.8
200	14.6	.717	104
300	17.7	.631	128
400	18.9	.592	156
500	20.2	.572	187
600	21.2	.567	222
700	22.0	.573	262
800	22.6	.585	308
900	23.1	.600	358
1000	23.6	.619	416
1500	25.3	.745	835
2000	26.7	.909	1622
*2500	27.6	0.082	3019
	$t=100°$		
1	.032969	$\bar{1}$.99900	0.99770
25	0.870	.975	23.6
50	1.86	.947	44.3
75	3.01	.921	62.5

TABLE VI.—*Continued.*

p ATMOS.	ρ MOL./LITER	LOG f/p	f ATMOS.
	$t=100°$—(*Continued*)		
100	4.37	0.894	78.3
150	7.69	.839	104
200	11.1	.790	123
300	15.2	.721	158
400	17.3	.683	193
500	18.7	.664	231
600	19.8	.658	273
700	20.6	.662	321
800	21.3	.674	378
900	21.9	.691	442
1000	22.4	.710	513
1500	24.4	.831	1018
2000	25.9	.987	1941
2500	27.0	0.1625	3634
	$t=125°$		
1	.030667	$\bar{1}$.99920	0.99816
25	0.803	.980	23.9
50	1.69	.959	45.5
75	2.67	.938	65.0
100	3.77	.918	82.8
150	6.26	.880	114
200	8.91	.843	139
300	13.1	.786	183
400	15.5	.754	227
500	17.1	.737	273
600	18.3	.734	326
700	19.3	.740	385
800	20.0	0.749	449
900	20.8	.763	521
1000	21.4	.777	599
1500	23.6	.887	1156
2000	25.0	0.0275	2131
2500	26.2	.180	3784
*3000	27.2	.338	6533
	$t=150°$		
1	.028843	$\bar{1}$.99932	0.99844
25	0.748	.983	24.0
50	1.55	.967	46.4
75	2.42	.951	67.0
100	3.36	.936	86.3
150	5.39	.906	121
200	7.54	.879	151
300	11.3	.835	205
400	13.9	.809	258
500	15.7	.797	314
600	17.0	.795	374
700	18.1	.798	440
800	19.0	.806	512
900	19.8	.820	595
1000	20.4	.838	689
1500	22.7	.952	1345
2000	24.4	0.0965	2498
2500	25.6	.253	4476
3000	26.7	.414	7783

(ii) Amagat densities into g/cc, divide them by 505.86, or multiply them by 0.0019768. (iii) Amagat densities into mol./l, divide them by 22.2623, or multiply them by 0.0449189.

4. THE PROCEDURE

In our previous work,[5] which had involved gases at temperatures considerably above their critical temperatures, we had smoothed and differentiated the data by means of large scale graphs of Δ and α plotted in isobars and isotherms against suitable functions of temperature or pressure. Δ was defined as $v(pv/RT-1)$ and α as $v(RT/pv-1)$. For the data on carbon dioxide, however, a deviation function of different behavior was desired: after some trials, the function

$$\sigma = Tv \log pv/RT, \qquad (40)$$

being one of several suggested to us by Mr. C. S. Cragoe of the National Bureau of Standards, was found to be satisfactory. Like the previous Δ and α, so also σ measures the departure from the perfect gas law, and magnifies the experimental irregularities, especially at low pressures.

The data presented by Michels are in uneven temperatures, e.g. 25.053°, 31.037°, ···, 75.260°, etc. Smoothing and reduction to integral temperatures and pressures was accomplished by drawing large scale plots of σ, one chart showing σ in isochores against t, another showing σ in isotherms against the density, ρ. These curves, after adjustment, represent traces of a σ, ρ, t surface. To obtain therefrom traces of the same surface at integral temperatures, values of σ were read from the smoothed isochores of σ *vs.* t at the desired integral temperatures (see Table VI). Having a smoothed value of σ at any chosen temperature and density, one can compute the corresponding pressure from the definition of σ, namely $Tv \log pv/RT$.

To find the densities at the integral pressures in Table VI, a chart was made showing σ *vs.* p in isotherms; σ read therefrom at any desired integral pressure establishes the corresponding

[5] W. Edwards Deming and Lola E. Shupe, I (nitrogen) Phys. Rev. **37**, 638–654 (1931); II (carbon monoxide) **38**, 2245–2264 (1931); III (hydrogen) **40**, 848–859 (1932); IV (entropies) **45**, 109–113; V (nitrogen) **48**, 448–9 (1935).

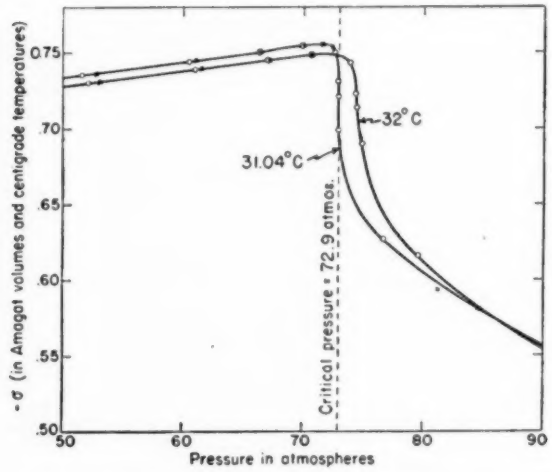


FIG. 29. Isotherms showing σ *vs.* p at 31.04° and 32° in the neighborhood of the critical region. The crosses are experimental points. The circles are interpolated from the σ *vs.* ρ isotherms (not shown) for assistance in drawing the curves. The σ *vs.* ρ isotherms exhibit no peculiarities in this region, being perfectly regular.

density (ρ or $1/v$) through the equation

$$\sigma = Tv \log pv/RT.$$

The σ *vs.* p isotherm at the critical temperature (31.04°C as given by Michels) exhibits a discontinuous slope at the critical pressure, as is seen in Fig. 29: this follows from the fact that at the critical point $(dp/dv)_T = 0$, wherefore $(d\sigma/dp)_T = \infty$. The vertical part of the 31.04° isotherm points to 72.9 atmos. as the critical pressure. The Michels give 72.85 atmos., which is in good agreement. Actually, a steep slope occurs also at temperatures just above the critical, as is seen in the isotherm for 32° in Fig. 29.

Fugacities were computed by graphical integrations according to the equation

$$\log f/p = -(0.4343/RT)\int_0^p \alpha\, dp, \qquad (28)$$

wherein $\alpha = v(RT/pv - 1)$, as previously defined. In place of graphical integrations, two terms of the Euler-Maclaurin summation

$$\int_1^{p_2} \alpha\, dp = \tfrac{1}{2}(\alpha_1+\alpha_2)(p_2-p_1) - \frac{1}{12}(p_2-p_1)^2\left[\frac{d\alpha}{dp}\right]_1^2 + - \cdots \qquad (41)$$

were useful when the α *vs.* p isotherms did not have high curvature. At high densities the integral in Eq. (41) was conveniently evaluated as $\int(\alpha v^2 dp/dv)_T d\rho$, for the reason that in this region the density scale is much compressed relative to the pressure scale.

5. THE RESULTS

The results are shown in Table VI. The calculations are considered to be consistent with the experimental data to as many figures as written. Extrapolations (never very long) are starred. Below the critical temperature the calculations are carried up to the saturation densities (marked by an *s* in the table), these points having been determined from the work of Myers and Van Dusen[6] on the vapor pressure of carbon dioxide.

The interpolation of fugacities for any intermediate argument (density, temperature, or pressure) is accomplished easier with $\log f/p$ than with f itself, and this is the reason for showing $\log f/p$. Moreover, the graphs of $\log f/p$ against p are nearly linear to 50 atmos.; then since $f=0$ and $\log f/p = 0$ at $p=\rho=0$, if one wished to know the fugacity (e.g.) at 3 atmos. and 40°, he would interpolate by proportional parts between $p=0$ and $p=25$, finding $\log f/p = (3/25)(\bar{1}.954) = \bar{1}.99448$, whence $f/p = 0.98737$, and $f = 3\times 0.98737 = 2.962$ atmos.

In conclusion it is a pleasure to mention the helpful interest of our friends Mr. C. S. Cragoe of the National Bureau of Standards, and Drs. A. and C. Michels in Amsterdam.

[6] C. H. Myers and M. S. Van Dusen, Nat. Bur. Stand. J. Research **10**, 381–412 (1933).

The Range and Validity of the Field Current Equation

F. R. ABBOTT AND JOSEPH E. HENDERSON
University of Washington, Seattle, Washington

(Received May 15, 1939)

Field currents from cold cathodes under high vacuum conditions were investigated over a range from 10^{-17} to 10^{-5} ampere. Electrometer measurements of currents below 10^{-10} ampere showed no deviation from the empirical equation of Millikan and Lauritsen. Precise measurements in the higher range exhibited a consistent deviation from the Millikan and Lauritsen equation, proportional to the fourth power of the field. This deviation is shown to be not necessarily in disagreement with the basis of the Fowler-Nordheim theoretical field current equation.

ONE of the early successes of wave mechanics was the development of an equation for cold cathode emission under high electrical fields, giving the first theoretical justification of the empirical equation

$$I = Ce^{-b/F} \tag{1}$$

obtained by Millikan and Lauritsen[1] in which I is the total emission current, b and c are constants, and F is the electric field intensity. In all cases where good vacuum conditions prevailed, the equation gave good agreement with experiment.[2] This investigation was undertaken to determine the range of validity of Eq. (1) and to search for possible deviations at high fields as predicted by wave-mechanical theory. The latter phase of the subject was briefly discussed by Stern, Gossling, and Fowler.[3]

The wave-mechanical solution of the problem by Oppenheimer, Fowler, and Nordheim[4] was based on the computation of the probability of penetration of the surface potential barrier by an electron incident upon the surface from within the metal. The simple form of surface barrier assumed by them leads to a transmission coefficient

$$D = \exp\left[-Bw^{\frac{3}{2}}/F\right], \tag{2}$$

where D is the probability of an electron of work function w, penetrating the surface barrier under the influence of an electric field F; and B is a numerical constant. It is apparent that theoretically any particular group of electrons of work function w would obey the empirical equation, (1). The total current is obtained from (2) by integration over the entire range of energies of the electrons and all values of the field over the emitting area.

Fowler and Nordheim assumed a Fermi distribution of energy and computed the emission per unit area under uniform field. A final assumption of an image force contribution to the surface potential barrier led to the theoretical field current equation

$$J = 1.55 \cdot 10^{-6}(F^2/w)10^{-2.98 \cdot 10^7 w^{\frac{3}{2}}\phi(y)/F}, \tag{3}$$

in which J = emission current in amperes per square centimeter; F = electric field in volts per centimeter; w = work function in electron volts; $y = 3.78 \cdot 10^{-4}(F/w)^{\frac{1}{2}}$ is the reduction in height of an image force potential barrier due to an impressed field F; ϕ is an elliptic function of y, the values of which have been computed by Nordheim, and yield the curve of Fig. 1.

In a strict sense, Eq. (3) applies only at 0° absolute. The temperature dependence deduced is expressed approximately by

$$J(T)/J(0) = 1 + (4 \cdot 10^8 T^2 w)/F^2, \tag{3a}$$

in which $J(T)$ and $J(0)$ are the specific emission currents at T°C and 0°C. With $F = 10^8$ volts per centimeter at room temperature,

$$J(T)/J(0) = 1.01.$$

At lower fields and higher temperatures the dependence increases slightly, but for logarithmic

[1] R. A. Millikan and C. C. Lauritsen, Proc. Nat. Acad. Sci. **14**, 45 (1928).
[2] R. A. Millikan and C. F. Eyring, Phys. Rev. **27**, 51 (1926). R. A. Millikan and S. S. Mackeowan, Phys. Rev. **31**, 900 (1928). General Electric Company of London Staff, Phil. Mag. **7** (1), 609 (1926).
[3] T. E. Stern, B. S. Gossling and R. H. Fowler, Proc. Roy. Soc. **A124**, 699 (1929).
[4] H. A. Bethe and A. Sommerfeld, *Handbuch der Physik* Vol. 24, Part 2, Sec. 3, Art. 19.

scales of J, as used experimentally, the temperature effect on field currents can ordinarily be disregarded. The almost complete temperature independence up to 1000° Kelvin is one of the outstanding characteristics of field emission.[5]

The usual check of the experimental field currents is to plot log I against the reciprocal of the applied voltage $1/V$. If Eq. (1) were of the correct form, a straight line would result since the field intensity F is proportional to the applied voltage.

Equation (3) in logarithmic form becomes

$$\log J = \log \frac{(1.55 \cdot 10^{-6})}{w} + \log F^2 - 2.98 \cdot 10^7 \frac{w^{\frac{1}{2}}}{F} \phi(y). \quad (3b)$$

For a constant emitting area under a uniform field the ratio of the observed total emission I to the emission per unit area J should be constant. If in Eq. (3b) the log F^2 could be disregarded and $\phi(y)$ remained unity, then log J would be linearly related to $1/F$ or, in measurable quantities, log I would be linearly related to $1/V$. However, the log F^2 and the factor $\phi(y)$ contribute toward a deviation from linearity in a log J, $1/F$ curve. In Fig. 2, log J, as computed from Eq. (3b), is plotted against $10^8/F$ for several values of the work function w. The deviation[6] of the $w=4.5$ curve from linearity as F increases from $1.5 \cdot 10^7$ to 10^8 is shown in Fig. 3. It is apparent that the deviation becomes increasingly pronounced as F approaches 10^8 volts per centimeter. The relative importance of the variable terms in Eq. (3b) is shown by Fig. 4. The main effect of the factor $\phi(y)$ is a vertical displacement of approximately two units of the straight line representing $-2.98 \cdot 10^7 (w^{\frac{1}{2}}/F)$ for $w=4.5$ and designated by $-b/F$. The two unit increase in log J would be equivalent to a hundred-fold increase in the emission J. The term, log F^2, from the figure deviates appreciably from a straight line function of the abscissa $10^8/F$.

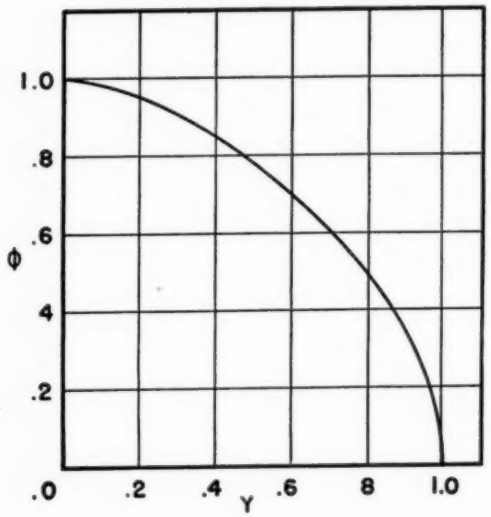


FIG. 1. Correction factor $\phi(y)$ introduced into the exponent of the Fowler-Nordheim equation, (3), to account for a reduction y in the height of an image force potential barrier due to an applied field F.

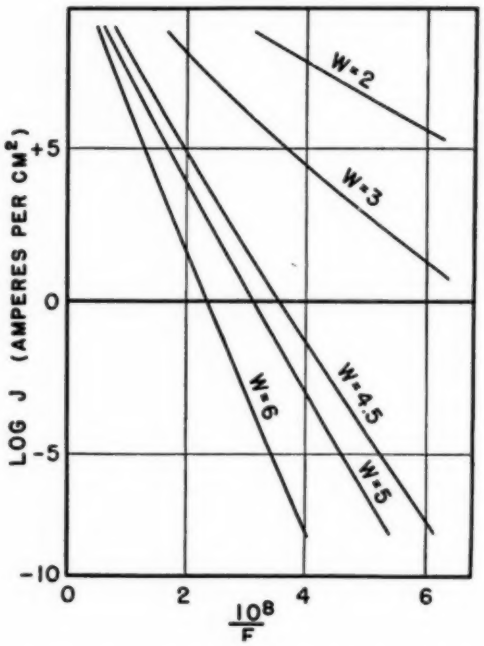


FIG. 2. The theoretical curves relating the emission to the field intensity for several values of the work function w.

Since experimentally the emitting area and the actual field intensity are difficult to even estimate, there appears no way to experimentally measure the effect of the ϕ factor. Measurements sufficiently precise should show evidence for or against the nonlinear log F^2 term, since F can be varied experimentally over a broad range.

[5] A. J. Ahearn, Phys. Rev. **44**, 277 (1933). Millikan and Eyring, reference 2.

[6] Deviations are the computed ordinate differences between the log J curve and a chord which closely approximates the log J curve in the vicinity of $10^8/F=3$.

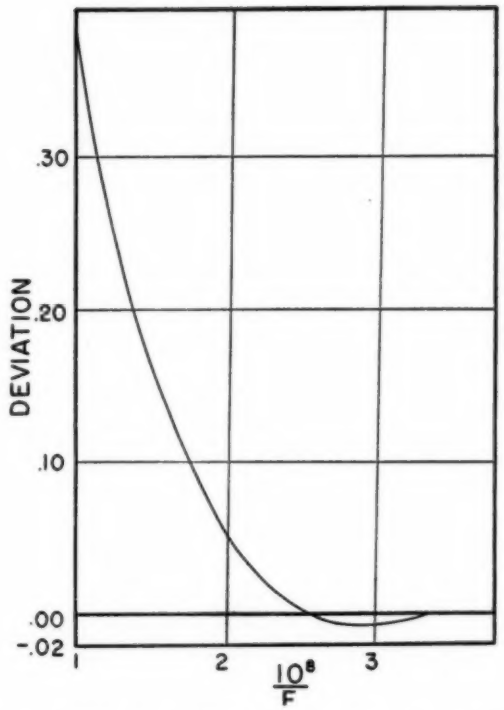


FIG. 3. Deviation of the $w=4.5$ curve of Fig. 2 from a straight line.

It is not difficult to produce a field which computed from the geometry approaches 10^7 volts per centimeter. Schottky[7] has shown that due to minute surface irregularities field intensity computed from the microscopic geometry may well be low by a factor of 10. Thus an experimental investigation of the deviation from linearity for geometrical fields of 10^6 and 10^7 volts per centimeter might yield information which would serve as a test of the theoretical Eq. (3).

The investigation to be described was divided into two parts. 1. By electrometer measurements of small emission currents and galvanometer measurements of larger currents from the same experimental tube, one curve of very broad range was obtained. 2. By the use of type K potentiometers a series of precise curves was obtained for the larger currents and higher field range.

EXPERIMENTAL PROCEDURE

Part I. Extended range field current measurements

The tube, constructed from a one-liter Pyrex flask employed a 0.001-centimeter tungsten filament and a one-centimeter diameter copper cylinder as anode. The entire internal assembly filament leads and their supporting inseal were surrounded by a grounded copper shield. The external filament circuit was completely shielded and evacuated along with the electrometer case for small current measurements.

The high vacuum system included a two-stage Kurth type mercury diffusion pump and a liquid-air trap immersed after thorough bake-out process. The metal parts were subsequently bombarded until, after cooling, consistent ionization gauge pressure indications were below $4\cdot10^{-8}$ mm of mercury.

Emission currents were measured by a FP54 pliotron electrometer circuit with floating grid. The modified Barth[8] type of circuit employed is shown by Fig. 5. For currents immediately below the galvanometer range a 0.001-microfarad radio type variable condenser with pressed amber insulation was introduced into the control grid circuit to reduce the rate of drift.

The calibration was checked at the junction with directly read galvanometer currents. Currents from 10^{-10} to 10^{-5} ampere were measured by a Leeds and Northrup Type HS galvanometer and Ayrton shunt.

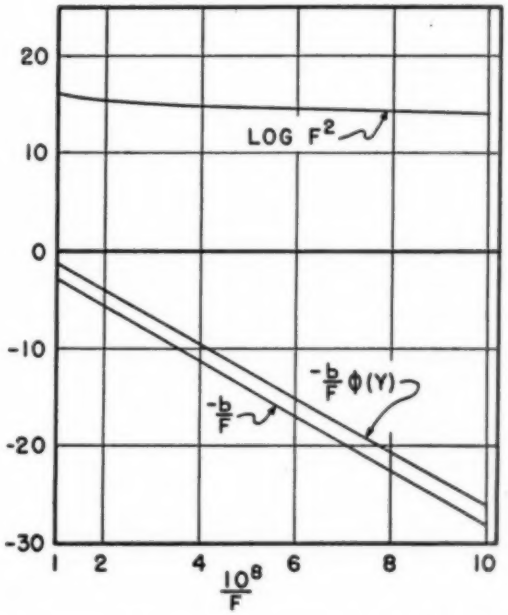


FIG. 4. Curves of the component terms in log J showing the relative magnitudes of log F^2, $-b/F$, and $-b\phi(y)/F$.

[7] W. Schottky, Zeits. f. Physik **14**, 63 (1923).

[8] O. B. Pennick, Rev. Sci. Inst. **6**, 115 (1935).

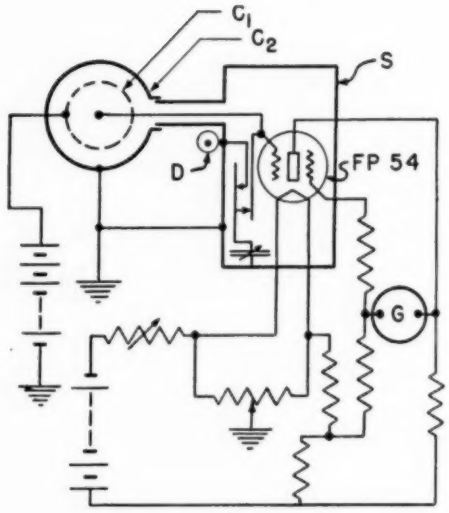


FIG. 5. Modified Barth type of FP54 Pliotron electrometer circuit used for small field current measurements. C_1—copper cylinder high voltage anode; C_2 and S—grounded electrostatic shielding system; D—control grid grounding device; G—Type H.S. galvanometer.

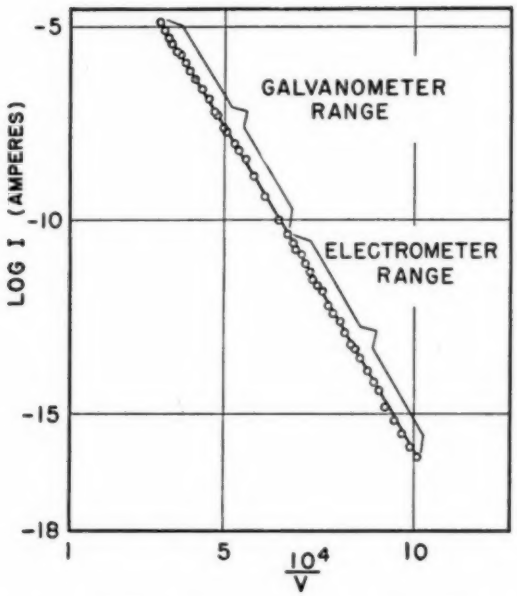


FIG. 6. Combined galvanometer and electrometer field current curve.

Throughout the electrometer range sufficiently steady high potentials were obtained from radio type B batteries. A small 3000-volt d.c. generator was used for larger currents and voltages of the galvanometer range.

Figure 6 shows the combined electrometer and galvanometer field current curve taken. Neither section exhibits curvature. No significance is attached to the slightly different slopes of the two sections. It was probably due to a partial breakdown of the emitting point as voltage was raised to the highest value at the beginning of the galvanometer run which was taken some time after the electrometer run.

The curve, extending over a current range of 10^{11} demonstrates the broad validity of the empirical Eq. (1). We conclude either that at the fields employed deviations do not exist, or else they are too minute to be observed with the instruments used.

Theory predicts increasing deviation with increased field, and this is the direction in which increased accuracy seemed relatively attainable by use of Type K potentiometer measurements. The foregoing work indicated that under best vacuum conditions the effect could be made steady enough to justify their use.

Part II. Precise field current measurements

A new tube was constructed, as shown in Fig. 7, to give higher fields at lower voltages. A 300-cc Pyrex flask was used for the tube itself. The anode was a carbon block and the cathode, a sharp tungsten point, was brought extremely close to the anode surface. A filament was provided for bombarding the electrodes. The evacuation process was essentially the same as that previously described. After bombarding, and while both electrodes were still considerably warmer than the glass walls, the cathode inseal was heated to the yield point and the tungsten point allowed to move into electrical contact with the carbon anode. Subsequent further cooling of the anode and cathode supports resulted in moving the point back a very small distance from the carbon. With this geometrical arrangement, galvanometer field currents were obtainable with less than 1000 volts applied. One objection to this method of producing high fields with low potentials is that the heating effect of currents approaching a milliampere produces changes in the anode cathode spacing due to expansion of the parts. This was in most cases avoided by opening the potential circuit for several minute intervals between the larger current readings.

Potentiometer readings of the drop across 20 ohms of an 885,000-ohm voltmeter resistor gave an accurate measure of the applied voltage.

Standard resistors of 1000 ohms to 1,000,000 ohms introduced in the cathode circuit were used in connection with a second Type K potentiometer to measure the field emission current.

Results and Discussion

The log I curve of Fig. 8 is typical of the results of six runs taken by this method with different separations of the electrodes. The actual form of the point was altered by honing three times during the course of the work. The double point near the top shows the reproducibility attainable for measurements taken at the beginning and end of a run. The curvature, although apparently slight, is evident when a straight edge is laid along the curve and is well beyond the probable error of measurement.

A comparison of the slopes of the theoretical curves of Fig. 2 with the experimental curve of Fig. 8 permits an estimate of the form factor θ relating applied voltage to the effective field ($F=\theta V$) for any assumed value of the work function w. From these form factors curves of log (I/F^2) as a function of $10^4/V$ were plotted. By theory these should yield straight lines if emission occurs from constant areas under uniform fields. This was done for assumed values of the work function from 3 to 9 electron volts. The

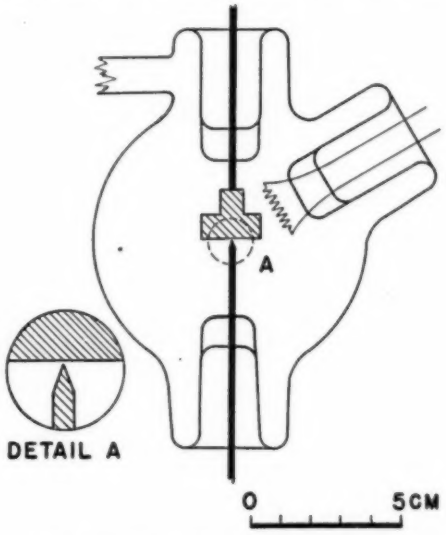


Fig. 7. Experimental tube for high field intensity at moderate applied potentials.

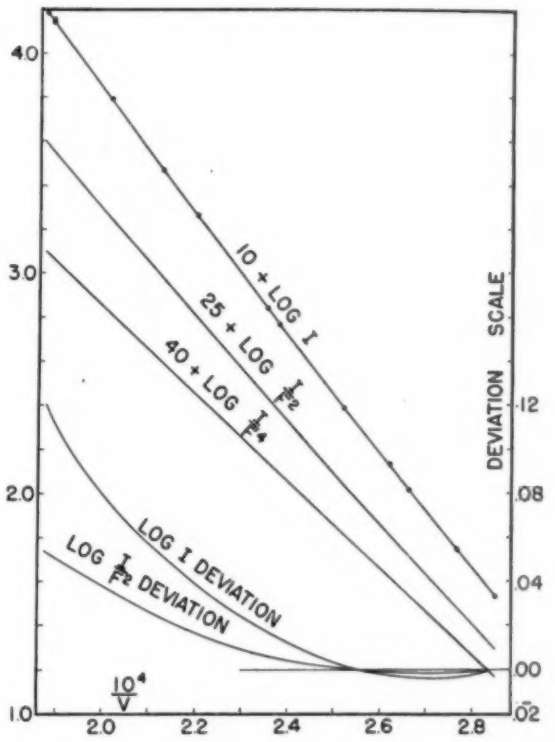


Fig. 8. Typical precise field current curves for the ordinates given adjacent to the curves. The double point at the top represents the reproducibility of the data; one point was taken at the beginning and the other at the end of the run. Deviations from linearity may be observed with a straight edge and are plotted at the bottom of the figure. The deviation of the $40 + \log (I/F^4)$ curve is less than the probable error of measurement and is not plotted.

log (I/F^2) curve for $w=4.5$ is shown in Fig. 8. A consistent deviation appeared to remain regardless of the assumed value of the work function. Further investigation indicated that experimental values of log (I/F^3) or log (I/F^4) would more closely approach linearity as a function of $10^4/V$. The log (I/F^4) is shown on Fig. 8. The deviation (computed as in Fig. 3) of the log I and log (I/F^2) curve are shown at the bottom of Fig. 8. The deviations of the log (I/F^4) curves are less than the probable error of measurement, and hence are not shown. We therefore conclude from this investigation that the functional dependence of the total field current emission upon the field is best represented by an equation of the form

$$I = CF^4 e^{-b/F}, \tag{4}$$

where C and b are constants.

Although the form of this equation is different from the Fowler-Nordheim equation, (3), it does

not necessarily contradict the basis upon which this equation is founded since this equation refers to the specific emission under constant area and uniform field. Consideration of this equation shows that any reasonable assumption as to the variation in field as a function of the distance from the point of maximum intensity leads to a higher power of F than occurs in the expression for J, the emission per unit area under constant field.

For example, in the case of an axially symmetrical point for which the field intensity at distance r from the axis is expressible in series form, the field in the vicinity of the axis is given by

$$F = F_0 + r(\partial F/\partial r)_{r=0} + r^2(\partial^2 F/\partial r^2)_{r=0} \cdots.$$

Since the axis is a maximum point $(\partial F/\partial r)_{r=0} = 0$ and thus

$$F = F_0 + r^2(\partial^2 F/\partial r^2)_{r=0}$$

and to first approximation putting

$$(\partial^2 F/\partial r^2)_{r=0} = -\alpha$$

we have

$$F = F_0 - \alpha r^2.$$

The total emission is given by

$$I = \int_0^\infty 2\pi r J(r) dr.$$

Introducing the above expression for F into the Fowler-Nordheim expression for J, Eq. (3), and integrating, we obtain approximately

$$I = \alpha F^3 e^{-\beta/F}, \tag{5}$$

in which all factors such as the work function w and the function ϕ have been included in constants α and β.

Only in the case of uniform field over the entire emitting area such as might result from a patch of low work function material covering a small area on the cathode would the Fowler-Nordheim expression, (3), describe the total emission. Under the more probable physical condition of an approximately symmetrical pointed electrode and constant work function over the area, Eq. (5) should more closely represent the total emission current. The experimental equation, (4), requires a still higher power of F. The correct power of F in the coefficient in any physical case depends on the exact microscopic shape of the point.

The cooperation of the Puget Sound Power and Light Company has been greatly appreciated throughout the entire work. The able laboratory assistance of Mr. Harold O. Wyckoff is gratefully acknowledged at this time.

LETTERS TO THE EDITOR

Prompt publication of brief reports of important discoveries in physics may be secured by addressing them to this department. Closing dates for this department are, for the first issue of the month, the eighteenth of the preceding month, for the second issue, the third of the month. Because of the late closing dates for the section no proof can be shown to authors. The Board of Editors does not hold itself responsible for the opinions expressed by the correspondents.

Communications should not in general exceed 600 words in length.

Electron Distribution in ZnO Crystals

In discussing the nature of the binding forces in various crystals one finds cases in which the usual criteria for their classification as electrostatic, homopolar, or polarization forces will not decide the point. We wish to report briefly on the application to ZnO of a method of approach to this problem through the interpretation of x-ray or electron diffraction intensities. A detailed account of this work will appear elsewhere.[1]

Intensities in the electron diffraction pattern from ZnO show[2] large deviations from values computed from the centrally symmetric charge distributions of Pauling and Sherman[3] for the Zn and O atoms. The anomalies are obvious on visual inspection, the order in intensity of quite prominent rings being reversed from that expected. Though x-ray intensities do not show these prominent anomalies, Lark-Horovitz and his co-workers were forced to the conclusion that they are due not to some factor in the experimental methods, but to the nature of the charge distribution in the crystal.

The ZnO crystal is of the hexagonal close packed type. Each zinc atom lies near the center of a tetrahedron formed by four oxygen neighbors, but nearer by 0.1A to one of these oxygen atoms than to the others; each pair of zinc-oxygen nearest neighbors lies parallel to the c axis. Yearian has noted that one might account qualitatively for the observed intensities by assuming a shift of the M shells of the zinc atoms along the c axis, toward the nearest oxygen neighbors. However, Johnson and Nordheim[1] have shown that any such polarization which might arise from intra-crystalline fields can account for only some 1/160 of the observed effect. We have thus been brought to the conclusion that the valence electrons in the crystal, despite their small numbers, are responsible for the observed anomalies.

Since a more direct test of this idea is not practicable at present, we have sought a simple and qualitatively reasonable model for the charge distribution in the crystal which will give the observed electron diffraction intensities. To the spherically symmetric distributions of Zn^{++} and O we have added various linear distributions of the valence electrons, or, with better results, distributions over the surfaces of ellipsoids of revolution. To account for the observed intensities we find it necessary to concentrate the valence electrons in the region between each zinc atom and its nearest oxygen neighbor; distributions which treat the four oxygen neighbors of a zinc atom as equivalent cannot give the desired results. Good agreement with Yearian's observations is given by a distribution of three ("valence") electrons over ellipsoids which enclose each pair of zinc-oxygen nearest neighbors. The major axis of the ellipsoids is 2.1A in length, the minor axis 0.35A. The greater part of the charge must be placed on the half of the ellipsoid nearer the zinc.

This model gives satisfactory agreement with observation for all strong lines, in particular accounting correctly for their order in intensity. Minor discrepancies remain, and the agreement is quite poor for the weak lines with the Miller index $l=4$. The accuracy of the experimental results does not justify a more elaborate attempt to reconstruct the charge distribution. However, we believe that the agreement obtained indicates conclusively that the actual electron distribution involves a concentration of valence electrons in the region between the zinc-oxygen nearest neighbors or a marked unidirectional distortion of the distribution of outer electrons in the zinc, such as is known to exist in those diatomic gases which can be treated theoretically.

K. Lark-Horovitz and C. H. Ehrhardt[4] have obtained further verification of this conclusion by carefully measuring x-ray diffraction intensities, from which they can deduce the corresponding electron intensities. Their observations show general agreement with the observed electron intensities, and particularly satisfactory agreement with intensities computed from a model like that described above, but with two electrons distributed over the ellipsoid instead of three.

The charge distribution thus found seems to indicate that in the crystal ZnO molecules tend to maintain their individuality, in contrast to the situation in purely ionic crystals. It suggests also that the homopolar binding of ZnO nearest neighbors is responsible in this crystal for the departure from the complete tetrahedral symmetry such as diamond exhibits. The ZnO pairs will of course not form an electrically symmetrical complex, and the binding of the molecules into the crystal will involve electrostatic and polarization forces, and probably some contribution from homopolar binding as well.

We are indebted to Drs. K. Lark-Horovitz and H. J. Yearian for many helpful discussions of this problem.

HUBERT M. JAMES
VIVIAN A. JOHNSON

Purdue University,
Lafayette, Indiana,
June 1, 1939.

[1] Johnson, Zeits. f. Krist., to appear soon.
[2] H. J. Yearian and K. Lark-Horovitz, Phys. Rev. **42**, 905 (1932). K. Lark-Horovitz, H. J. Yearian and J. D. Howe, Phys. Rev. **47**, 331 (1935). H. J. Yearian, Phys. Rev. **48**, 631 (1936).
[3] L. Pauling and J. Sherman, Zeits. f. Krist. **81**, 1 (1932).
[4] K. Lark-Horovitz and C. H. Ehrhardt, Phys. Rev. **55**, 605 (1939).

Super-Novae and the Neutron-Core Stars

The so-called super-novae, according to W. Baade and F. Zwicky,[1] are stars of visible radiation at maximum, per second, of the order of magnitude of 6.3×10^7 times the radiation of the sun, i.e., $Lv=2.38\times10^{41}$ erg/sec. It seems certain that, as assumed by Baade and Zwicky, a super-nova represents a star in a process of explosion. Consequently they postulate that a super-nova must emit particles of very high energy. As they do not suppose any concrete physical model to represent this type of stellar explosion, it is not possible to guess what is the ratio of the visible radiation to the total radiation (radiation plus emission of particles). Therefore their theoretical considerations, from thermodynamic relations deduced for pure radiation at equilibrium, are inapplicable to the case of a stellar explosion with emission of particles.

In another paper Baade and Zwicky[2] advance the suggestion that a super-nova represents the transition of an ordinary star into a neutron star, consisting mainly of neutrons.

Let us analyze this last suggestion. The stability of a neutron core in a star has been studied by Landau.[3] To transform, e.g., an oxygen nucleus into 16 neutrons, an energy of 0.008 mass unit per particle or 1.2×10^{-5} erg is required. Landau has shown that if the core has a uniform density of the order of 10^{14} g/cm^3, matter would be transformed into neutrons in the stellar interior, only if a neutron core of mass $M>10^{32}$ g$=0.05\odot$ already exists, where $\odot$ is the mass of the sun. If the mass of the neutron core is smaller than the given limiting value, the core would be transformed by an explosive process into different nuclei. An ordinary star is a gaseous star without neutron core. Therefore, for the reason given before, it is very difficult to imagine how a super-nova could result from the transformation of an ordinary star into a neutron star. Moreover, it is very simple to see that the process of formation of a neutron core can never produce an explosion as required to explain the appearance of a super-nova. In fact, if the neutron core has already a mass greater than the value given by Landau, the transformation of nuclei into neutrons would be an exothermic process and the star would emit energy. If the mechanism of the star cannot get rid of all the energy produced in this process, the temperature in the stellar interior would increase and consequently the mean density of the neutron core would decrease, which would produce a decrease in the gain of gravitational energy per gram of matter transformed. This effect reduces the velocity of reaction of the process, instead of increasing it, as would be necessary in Baade and Zwicky's theory. From what we have already stated, it may be imagined that the direction of the process may change, producing evaporation of neutrons in an adiabatic process and reducing again the temperature in the stellar interior. Then the process of neutron core formation could start again. This mechanism of contraction and expansion perhaps occurs in some variable stars.

We would like to point out that in Landau's theory the gravitational energy of the gaseous matter in the star is neglected in comparison with the gravitational energy of the neutron core per gram of matter. This assumption is good enough when the radius of the star is much larger than the radius of the neutron core, but it is not true when the mass of the neutron core is of the same order of magnitude as the total mass of the star. When the neutron core increases beyond a certain limit, it is not possible to neglect the gravitational energy of the gaseous layer; this effect also tends to stop the growth of the neutron core. Consequently we may conclude that the formation of a neutron core in a stellar interior cannot produce an exothermic explosive reaction.

In a paper to be published elsewhere we analyze the possible existence of a neutron core taking into account the existence of short range interaction forces among the neutrons in the neutron core. We prove that in this case the neutron core can be split into two different phases: a gaseous and a condensed phase. In this way we show that the formation of a neutron core can provide a concrete physical mechanism for the collapsing stars, and that this process cannot produce any explosion.

I should like to express my thanks to Professor M. S. Vallarta for useful discussions on this subject.

FÉLIX CERNUSCHI*

Massachusetts Institute of Technology,
Cambridge, Massachusetts,
June 12, 1939.

* Fellow of the Argentine Association for the Progress of Science.

[1] W. Baade and F. Zwicky, Proc. Nat. Acad. Sci. **20**, 254 (1934).
[2] W. Baade and F. Zwicky, Proc. Nat. Acad. Sci. **20**, 259 (1934).
[3] L. Landau, Nature **141**, 333 (1938).

A Tentative Theory of the Origin of Cosmic Rays

Since the publication of Bethe's paper on energy production in stars[1] it is very difficult to understand how the huge amount of energy produced by a super-nova[2] could be explained in terms of the known nuclear reactions which might take place in a stellar interior. In the following lines we propose a new mechanism which can provide enough energy both to explain the appearance of a super-nova and the origin of cosmic radiation.

Let us assume that the white dwarfs are stars with a neutron core the density of which increases in the process of cooling. When the density of the neutron core is of the same order as the density of matter in nuclei we cannot be sure that the neutron core is going to be stable. In fact, if we assume that the binding energy between proton-neutron is at least as high as the binding energy between neutron-neutron, a non-negligible amount of neutrons will be transformed into protons, thus reducing the free energy of the system. The conservation of the electrical charge then forces the creation of an equal number of electrons.

In this way it is possible to imagine, bearing in mind the high pressure in the interior of the core, the formation of unstable nuclei of very high atomic number.

As is well known, the nuclei of uranium and thorium[3] if bombarded by neutrons split up into two nuclei of nearly the same mass with a release of energy of the order of 100 Mev per secondary nucleus. This process of fission could take place with the unstable nuclei of higher atomic

number whose existence in the neutron core we have postulated.

On these assumptions it is easily seen that the fission of a nucleus of atomic number 10,000 would produce energies of the order of 10^{12} ev.

According to our model a super-nova would not be a transition of an ordinary star into a neutron star, but would result from the explosion of the neutron core of a white dwarf. According to our model it is possible to imagine a process of successive fissions; i.e., a nucleus of very high atomic number would be split up into two nuclei each of which would be split up again into another two nuclei and so on. In this way we should have a mechanism which would somewhat increase the energy of the particle in each fission.

To understand the possibility of producing in the neutron core nuclei of very high atomic number, let us suppose that all the different types of short range forces, on the average, produce a potential energy, for the whole nucleus, proportional to $-N^2\epsilon_1$, where N is the atomic weight; we also assume that the repulsive forces due to the charged particles produce a potential energy proportional to $+Z^2\epsilon_2$, where Z is the atomic number. Therefore, in this crude model for nuclei of very high atomic number, we should have the following condition:

$$Z^2\epsilon_2 - N^2\epsilon_1 \lesssim 0. \tag{1}$$

We may calculate roughly the ratio ϵ_2/ϵ_1 supposing that for the uranium nucleus (1) is zero; this yields

$$N^2/Z^2 = \epsilon_2/\epsilon_1 \eqsim 7. \tag{2}$$

Thus, if we take $Z = 10{,}000$ we obtain $N \eqsim 26{,}000$. Assuming a uniform density for matter in the interior of nuclei, and bearing in mind that the range of nuclear forces is smaller than 10^{-12} cm, we may write the following relation which limits the radius of super-nuclei:

$$(x/239)^{\frac{1}{3}} \lesssim 10 \tag{3}$$

where x is the atomic weight of the super-nucleus. If x is 26,000 the ratio (3) is 5. Because of the range of nuclear forces, this value is an upper limit; and, consequently, the limiting value of the atomic weight of super-nuclei will be of the order of 10^4. As we have seen, the fission of these super-nuclei can produce energy of the order of 10^{12} ev, which should be the highest value of the energy of the cosmic rays. Under these assumptions it is possible to imagine a concrete physical mechanism which might underlie the production of cosmic rays. We might also point out that the present theory also leads one to expect multiple charged primary particles in cosmic radiation, as already suggested on other grounds.[3]

I should like to express my thanks to Professor M. S. Vallarta for useful discussions on this subject.

FÉLIX CERNUSCHI*

Massachusetts Institute of Technology,
Cambridge, Massachusetts,
June 12, 1939.

* Fellow of the Argentine Association for the Progress of Science.

[1] H. A. Bethe, Phys. Rev. **55**, 434 (1939).
[2] F. Cernuschi, "Super-Novae and the Neutron-Core Stars." Previous communication.
[3] M. S. Vallarta, Phys. Rev. **55**, 583 (1939).

Heat Capacity of Potassium Dihydrogen Phosphate at the Curie Point*

The anomalous heat capacity at the Curie point of a ferroelectric salt is related to the spontaneous polarization by the equation

$$\Delta C = -\tfrac{1}{2} f \cdot \partial P^2/\partial t \text{ erg/cm}^3\,^\circ\text{C},$$

where f, the Lorentz factor, has values of the order of magnitude of $4\pi/3$. This result,as in the corresponding case for ferromagnetics, is based on the inner field theory of Weiss. Recent investigations of the heat capacity of Rochelle salt by Hicks and Hooley[1] and Wilson[2] have shown that any anomaly at the Curie points of this salt must be less than one percent, in agreement with theory.

Professor Hans Mueller called our attention to the fact that, on the basis of the above equation and the dielectric data of Busch,[3] KH_2PO_4 should have a large anomaly at the Curie point, and at his suggestion we have measured the heat capacity of this substance. The results, as shown in Fig. 1, give a curve of heat capacity *versus* temperature

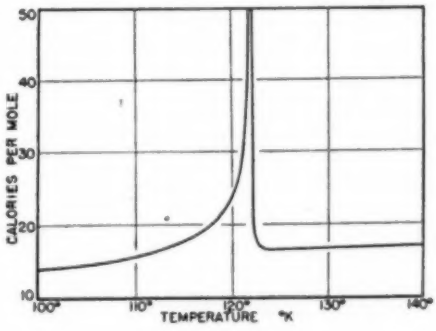


FIG. 1. Heat capacity of KH_2PO_4 as a function of temperature.

which agrees qualitatively with the predictions of the theory. The measurements, made in the calorimeter used by Hicks and Hooley, show that the maximum in the heat capacity occurs at 122°K. Busch gives 115°K for the Curie temperature, but his temperature scale is probably unreliable.

The rise of the heat capacity is of the correct order of magnitude. From the maximum spontaneous polarization given by Busch and the area under the anomalous portion of the heat capacity curve, the value $f = 0.7$ is calculated.

The present measurements cover the region from 80°K to 300°K. The measurements will be extended to 15°K, and a complete report of the investigation will be published later.

C. C. STEPHENSON
J. G. HOOLEY**

Research Laboratory of Physical Chemistry,
Massachusetts Institute of Technology,
Cambridge, Massachusetts,
June 15, 1939.

* Contribution from the Research Laboratory of Physical Chemistry, Massachusetts Institute of Technology, No. 425.
** Royal Society of Canada Fellow.

[1] Hicks and Hooley, J. Am. Chem. Soc. **60**, 2994 (1938).
[2] A. J. C. Wilson, Phys. Rev. **54**, 1103 (1938).
[3] G. Busch, Helv. Phys. Acta **11**, 269 (1938).

The Velocity of Sound in Liquid Helium under Pressure

The authors[1] have previously measured the velocity of sound in liquid helium evaporating under its own vapor pressure, using a resonance method with a quartz oscillator operating at 1338 kc. In these experiments no discontinuity was observed at the λ-point, whereas Ehrenfest's[2] thermodynamic relations for a change of state in which there is no latent heat require a sudden change in the compressibility, amounting to

$$\Delta K_T = -\Delta\alpha/(dp/dT)_\lambda,$$

where $\Delta\alpha$ is the change in the coefficient of expansion. It was possible that the measurements did not give the true velocity in the liquid, owing to the formation of minute bubbles in the sound field, although the liquid helium was not allowed to boil while measurements were in progress, and the hydrostatic pressure was considerably greater than the pressure amplitude of the waves.

The experiment has now been repeated with the liquid helium under an external pressure of 1 to 5 atmospheres. For this purpose the oscillator and reflector were enclosed in a copper bomb inside the cryostat proper. With boiling liquid helium surrounding the bomb in order to maintain any desired temperature, the liquid helium inside the bomb could be subjected to pressure by means of helium gas from a cylinder. Measurements of the velocity made under pressure in the neighborhood of the λ-point are shown in Fig. 1, along with the curve of the previous measurements.

Table I below shows the change ΔK_T in the isothermal compressibility, as calculated from the values of $\Delta\alpha$ and $(dp/dT)_\lambda$ estimated from the data of Keesom and Miss Keesom.[3] The last two columns give the approximate observed change ΔW in the velocity of sound, and the derived change ΔK_φ in the adiabatic compressibility. The calculated ΔK_T, and the experimental ΔK_φ are comparable to within one percent, since the ratio of the specific heats is estimated to be less than 1.01 in both He I and He II near the λ-transition.

It is hoped to continue this work far enough to obtain experimental values of the compressibility throughout the region of pressures less than about 5 atmospheres, where it is difficult to deduce values accurately from the existing pressure-volume data.

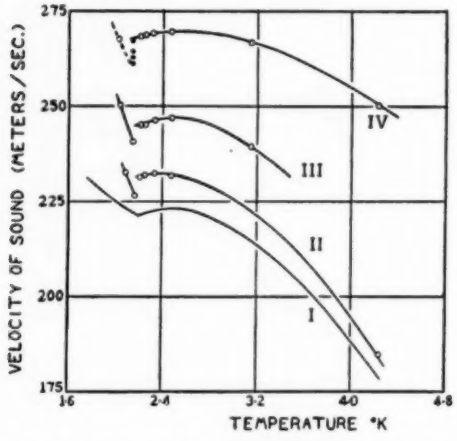


FIG. 1. The velocity of sound in liquid helium under various pressures, I vapor pressure, II 1 atmos., III 2.47 atmos., IV 5.55 atmos.

TABLE I. *Change in the compressibility of liquid helium at the λ-transition for various pressures.*

PRESSURE (ATMOS.)	VAP. PRESS.	1.0	2.47	5.55
$\Delta\alpha$	0.0499	0.036	0.039	0.039
$(dp/dT)_\lambda$ (c.g.s. 10^{-6})	−82	−87	−84	−79
ΔK_T calc. (c.g.s. 10^{10})	− 6.2	− 4	− 4.7	− 4.8
ΔW (m-sec.$^{-1}$)	0	6	5	7
ΔK_φ exp. (c.g.s. 10^{10})	0	− 7	− 5	− 5

J. C. FINDLAY
A. PITT
H. GRAYSON SMITH
J. O. WILHELM

McLennan Laboratory,
University of Toronto,
Toronto, Canada,
June 7, 1939.

[1] J. C. Findlay, A. Pitt, H. Grayson Smith and J. O. Wilhelm, Phys. Rev. **54**, 506 (1938).
[2] P. Ehrenfest, Proc. Kon. Akad. Amsterdam **36**, 153 (1933) (Leiden Comm. Suppl. 75b).
[3] W. H. Keesom and Miss A. P. Keesom, Proc. Kon. Akad. Amsterdam **36**, 482 and 612 (1933), Physica **1**, 128 (1933) (Leiden Comm. 224d, e, Suppl. 76b).

The Multiple Scattering of Charged Particles

The total scattering cross section for a Coulomb interaction is infinite. This fact has been a serious difficulty in the interpretation and calculation of scattering phenomena for charged particles. Attempts have been made to avoid the divergence of the formulas by observing that the field of a nucleus does not extend beyond a certain range because of the screening effect of the orbital electrons. The calculated results, however, have proved to be very sensitive to the choice of the distance and the manner in which the Coulomb field is cut off.

A further analysis of the divergence of the scattering formulas indicates that the "long range" of the Coulomb field is only part of the cause. Another reason is that multiple scattering has not been considered adequately. In the case of long range forces appreciable scattering due to many small deflections is quite prevalent. If this is properly taken into account, it is seen that the final results are no longer so sensitive to the cut-off radius of the Coulomb field. A method for treating this multiple scattering problem is given in a paper by Ornstein[1] which deals with the diffusion of neutrons in water. In this paper expressions are given for the angular distribution after ν collisions for any law of interaction. Combining these results with the well-known formulas for the probability that ν collisions will occur, one obtains a final scattering distribution.

This method, unfortunately, does not yield an explicit distribution function for the scattering angle θ, but gives it in terms of averages of $\cos\theta$, $\cos^2\theta$, etc., or of Legendre polynomials.

For the Rutherford scattering of electrons by a foil of

N atoms per cm², we find, for example, the following averages

$$\cos\theta = \left(1+\frac{\rho^2}{k^2}\right)^{-2\pi k^2 N},$$

$$P_2(\cos\theta) = \left(1+\frac{\rho^2}{k^2}\right)^{-6\pi k^2 N e^{+6\pi k^2 N}}.$$

In these expressions ρ is the cut-off radius of the Coulomb field and $k^2 = Z^2e^4(1-\beta^2)/m^2c^4\beta^4$. The exponent is usually quite small and hence the result is very insensitive to a change in ρ. For example, for 200 kev electrons passing through 1 mil of Al, we have $k = 5.4\times10^{-12}$ and $N = 1.5\times10^{20}$. This gives for the angle θ corresponding to the average cosine a variation of only from 41° to 47° while ρ changes by a factor of 10 from 10^{-9} to 10^{-8} cm. For 1 Mev electrons the corresponding variation is from 12° to 14° for the same range of ρ.

A very rough approximation for small angles gives the same Gaussian distribution as the diffusion treatment of multiple scattering of Bothe.[2] A slightly better approximation for small angles is

$$F(\theta)d\theta = 2\pi\sin\theta f(\theta)d\theta = \tfrac{1}{2}\lambda\exp(-\lambda\sin^2\theta/2)\sin\theta d\theta,$$

where

$$1/\lambda = 2\pi k^2 N\log\rho/k.$$

Further details of the calculations and results will be published later. We acknowledge the support of this work provided by the Horace H. Rackham Fund.

S. GOUDSMIT
J. L. SAUNDERSON

University of Michigan,
Department of Physics,
Ann Arbor, Michigan,
June 6, 1939.

[1] L. S. Ornstein, Proc. Amsterdam **XL**, 464 (1937).
[2] W. Bothe, Zeits. f. Physik **4**, 300 (1921).

The Existence of Radioactive Al^{29}

The bombardment of magnesium with α-particles produces two well-known radioactive bodies. They are Al^{28} and Si^{27} produced by the reactions Mg^{25} (α, p) Al^{28} and Mg^{24} (α, n) Si^{27}. The half-lives are 2.3 minutes[1] and 6.7 minutes,[2] respectively. In addition to these two it has been reported[1] that a third radioactive isotope is produced. It was assumed to be Al^{29} formed from the third magnesium isotope by the reaction Mg^{26} (α, p) Al^{29}. The activity produced with an α-particle source of 100 millicuries of radon was not sufficient for analysis and the half-life was estimated as 11 minutes.

The same sample has now been bombarded with the 16 million volt α-particles from the cyclotron and with the increased intensity the long period is found to be 4.0 hours. However, another sample of very pure magnesium supplied by the Dow Chemical Company does not show this activity. The decay curve found with the pure sample is shown in Fig. 1. It can be entirely explained by the existence of two radioactive bodies. The curve gives a half-life at 6.4 ± 0.1 minutes for Si^{27} and by subtraction the half-life of Al^{28} is found to be 2.3 minutes.

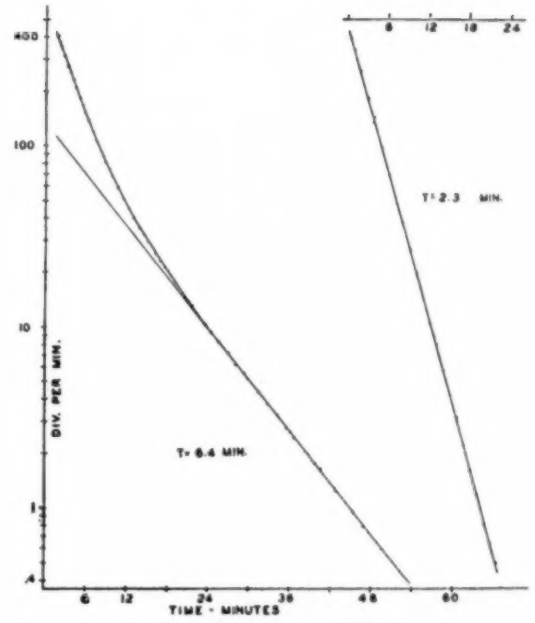


FIG. 1. Decay of magnesium +alpha-particles.

The third activity observed on the original sample must be due to an impurity. A likely impurity in magnesium is calcium and it is known that α-particles produce from calcium radioactive Sc^{43} with a half-life of 4 hours. With a pure calcium target the 4.0-hour activity is found to be about 300 times stronger than that observed with the impure magnesium sample. An impurity of 0.33 percent of calcium in the target would therefore explain our results. We conclude that the reaction Mg^{26} (α, p) Al^{29} does not occur to any appreciable extent.

W. J. HENDERSON
R. L. DORAN

Purdue University,
Lafayette, Indiana,
June 1, 1939.

[1] Ellis and Henderson, Proc. Roy. Soc. **156**, 358 (1936).
[2] Fahlenbrach, Zeits. f. Physik **96**, 503 (1936).

Transitions Between Glow and Arc Discharge

The experiments on this subject recently described by Mr. Fan[1] were performed under conditions such that high frequency glow-arc transitions might in some cases have been expected. Just thirty years ago G. W. Vinal and I[2] described transitions somewhat similar to Fan's, using various metals as electrodes, in air or nitrogen. Discontinuous changes from one type of discharge to the other occurred every second or so, the intervals being considerably shorter than in Fan's experiments. But we found that at currents of about 0.1 amp., and with very short gaps, several thousand transitions occurred each second. In illuminating gas or a mixture of hydrogen and acetone vapor, with a d.c. supply voltage from 400 to 1000, and with a circuit having no large inductances or resistances

within several meters of the discharge tube, the transitions took place so rapidly as to generate oscillations of a frequency of the order of 250,000. As much as 66 percent of the d.c. energy was converted into these high frequency oscillations. Several rather paradoxical demonstrations are described in the papers referred to above. For example, a miniature incandescent lamp in series with the discharge tube glows more brightly when the discharge is taking place than when the electrodes are short-circuited, although in the latter case a d.c. ammeter indicates a larger current.

The effect was independent of the material of the anode, and could be obtained to a greater or less extent with cathodes of metals ranging all the way from soft solder to Ag or Pt, and even with carbon. Oscillations ceased with currents above 0.8 amp.

With an a.c. 60-cycle supply the effect could also be observed, although not so well. In this connection attention may be called to a paper by Plesse[3] who found what appears to be the same high frequency glow-arc transitions with an a.c. supply.

WALTER G. CADY

Scott Laboratory,
Wesleyan University,
Middletown, Connecticut,
June 1, 1939.

[1] H. Y. Fan, Phys. Rev. **55**, 769 (1939).
[2] W. G. Cady and G. W. Vinal, Am. J. Sci. **28**, 89 (1909) and Physik. Zeits. **10**, 569 (1909); W. G. Cady, Am. J. Sci. **28**, 239 (1909) and Physik. Zeits. **10**, 623 (1909). This work is discussed in the book by K. W. Wagner, *Der Lichtbogen als Wechselstromerzeuger* (Leipzig, 1910).
[3] H. Plesse, Ann. d. Physik, **22**, 473 (1935). A somewhat similar phenomenon in the spark discharge is recorded by H. Kaiser and A. Wallraff, Zeits. f. Physik **112**, 215 (1939).

Initial Performance of the 60-Inch Cyclotron of the William H. Crocker Radiation Laboratory, University of California

During the past few weeks, we have been engaged in the adjustments of the 60-inch cyclotron of the William H. Crocker Radiation Laboratory, and at this time we wish to report its initial performance.

As is always fruitful in first turning on a cyclotron, we looked for resonance in the first instance by placing a Geiger counter nearby. With hydrogen, proton resonance was found close to the expected value of magnet current, and we proceeded to build up the intensity of the resonance effect by adjusting the magnetic field with shims. Following this adjustment, the beam to the target was observed and further shimming yielded 25 microamperes at eight-million volt protons. Probe measurements indicated about 100 microamperes circulating within the dees.

Next the hydrogen was replaced with deuterium at a pressure about one-tenth that normally used in our 37-inch cyclotron, and the procedure of adjustment was repeated. Again resonance of deuterons was first observed with a Geiger counter and shims were used to build up the effect. Indeed, the radiation intensity almost immediately obtained in this way, even with but one-tenth of the normal deuterium pressure, exceeded that ever obtained from the 37-inch cyclotron. At this juncture the high energy deuteron beam was looked for by detecting ionization in the air outside an aluminum target window. The beam was immediately found, and adjustments proceeded until ten microamperes to the target itself was obtained. The deuterons were observed to emerge from the target window and to penetrate the air for a distance of more than 1½ meters, indicating the energy to be in the neighborhood of 16 million volts, a value consistent with the frequency and the geometry. Under the circumstances, as is well known, replacing the deuterium with helium would result in a beam of 38 million volt alpha-particles. Dependence of the deuteron beam on the adjustments shows that very large currents are obtainable, perhaps as large as can conveniently be used. It is perhaps noteworthy also that this very high energy deuteron beam is obtained with an oscillator input of only 60 kilowatts. We therefore see no difficulties in the way of producing with the present equipment 25 million volt deuterons and 50 million volt alpha-particles, and moreover we are convinced that much higher energies could be obtained from a cyclotron of larger dimensions.

We wish to take this opportunity to express our profound gratitude to many individuals and organizations who have made these developments possible. In the first place we are indebted to the late Mr. Francis P. Garvan and to Mr. William H. Buffum of the Chemical Foundation, who provided the initial funds for the cyclotron construction, and to the late Mr. William H. Crocker, who provided for the laboratory building. We are equally indebted to the Rockefeller Foundation and the National Advisory Cancer Council for continued support which has made possible the rapid development of the work. We also are deeply appreciative of the important part our laboratory colleagues are having in this work, for all of our associates have been actively interested and have contributed in large measure to the development of the 60-inch cyclotron; and we are likewise appreciative of the invaluable help of the staff of the shop.

ERNEST O. LAWRENCE
LUIS W. ALVAREZ
WILLIAM M. BROBECK
DONALD COOKSEY
DALE R. CORSON
EDWIN M. MCMILLAN
W. W. SALISBURY
ROBERT L. THORNTON

William H. Crocker Radiation Laboratory,
University of California,
Berkeley, California,
June 12, 1939.

THE AMERICAN INSTITUTE OF PHYSICS

INCORPORATED

ADMINISTRATIVE STAFF

HENRY A. BARTON, *Director*

MADELINE M. MITCHELL, *Publications Manager*

Office: 175 Fifth Avenue, New York, New York

JOURNALS

THE PHYSICAL REVIEW
REVIEWS OF MODERN PHYSICS
JOHN T. TATE, *Editor*
J. W. BUCHTA, *Assistant Editor*

JOURNAL OF APPLIED PHYSICS
ELMER HUTCHISSON, *Editor*

THE JOURNAL OF THE OPTICAL SOCIETY OF AMERICA
THE REVIEW OF SCIENTIFIC INSTRUMENTS
F. K. RICHTMYER, *Editor*

JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA
F. R. WATSON, *Editor*

THE JOURNAL OF CHEMICAL PHYSICS
HAROLD C. UREY, *Editor*
JOSEPH E. MAYER, *Assistant Editor*

THE AMERICAN PHYSICS TEACHER
DUANE ROLLER, *Editor*

GOVERNING BOARD

For The American Physical Society:

JOHN T. TATE, *Chairman*
University of Minnesota

GEORGE B. PEGRAM, *Secretary*
Columbia University

KARL T. COMPTON
Massachusetts Institute of Technology

For The Optical Society of America:

E. C. CRITTENDEN
National Bureau of Standards

A. C. HARDY
Massachusetts Institute of Technology

W. B. RAYTON
Bausch & Lomb Optical Company

For The Acoustical Society of America:

FLOYD A. FIRESTONE
University of Michigan

HARVEY FLETCHER
Bell Telephone Laboratories

WALLACE WATERFALL
The Celotex Corporation

For The Society of Rheology:

WHEELER P. DAVEY
Pennsylvania State College

MELVIN MOONEY
United States Rubber Products, Inc.

R. L. PEEK, JR.
Bell Telephone Laboratories

For The American Association of Physics Teachers:

PAUL E. KLOPSTEG
Central Scientific Company

FREDERIC PALMER, JR.
Haverford College

F. K. RICHTMYER
Cornell University

The American Institute of Physics is a membership corporation of which five societies are the founding members. These five societies are the principal national societies devoted to pure and applied physics in America. Their names, and those of their representatives who form the Governing Board of the Institute, are given above. The principal work of the Institute is the publication of journals sponsored by the societies or by itself. The names and the editors of the journals are also listed above.

The broad purpose of the Institute is to represent, in all matters of wide or common interest, the five thousand or more members and subscribers associated with the Founder Societies and the journals. It aims to advance the science and profession of physics, and to promote cooperation between pure research, the applied sciences and the industries. To achieve these aims, it is greatly in need of an endowment and is empowered by law to administer grants and funds of any kind.

CONTENTS

PHYS. REV. 56 (1939) JULY 1

Investigation of the Products of the Disintegration of Uranium by Neutrons. PHILIP ABELSON 1

Intensity of Neutrons of Thermal Energy in the Atmosphere at Sea Level . . . . . C. G. MONTGOMERY AND D. D. MONTGOMERY 10

Self-Diffusion of Copper . . . . . J. STEIGMAN, W. SHOCKLEY AND F. C. NIX 13

Natural Radioactivity of Lutecium . . . . . W. F. LIBBY 21

Precise Determination of the Energy Released in the Production of Deuterium from Beryllium under Proton Bombardment. . . . . LESTER S. SKAGGS 24

Atomic Beam Study of the Hyperfine Structure of the Metastable $^3P_{2,3}$ State of In^{115} I. The Electric Quadrupole Moment of In^{115} . . . . . DONALD R. HAMILTON 30

Stark-Effect Broadening of Hydrogen Lines. II. Observable Profiles . . LYMAN SPITZER, JR. 39

Zeeman Effect of the Hyperfine Structure of Optically Excited Mercury Resonance Radiation . . . . . E. HOBART COLLINS 48

Paschen-Back Effect. V. Theory of the Effect for Intermediate Coupling . . . . . J. B. GREEN AND J. F. EICHELBERGER 51

Paschen-Back Effect. VI. The Spectrum of Neon. . J. B. GREEN AND J. A. PEOPLES, JR. 54

Electron Diffraction Studies of Thin Films. I. Structure of Very Thin Films . L. H. GERMER 58

On the Self-Energy and the Electromagnetic Field of the Electron. . . V. F. WEISSKOPF 72

Dispenser Cathode. A New Type of Thermionic Cathode for Gaseous Discharge Tubes . . . . . A. W. HULL 86

Change of Resistance in a Magnetic Field. . . . . LEVERETT DAVIS, JR. 93

Diffraction Theory of Electromagnetic Waves. . . . . J. A. STRATTON AND L. J. CHU 99

Some Physical Properties of Compressed Gases. VI. The Fugacity of Carbon Dioxide. . . . . W. EDWARDS DEMING AND LOLA S. DEMING 108

Range and Validity of the Field Current Equation. F. R. ABBOTT AND JOSEPH E. HENDERSON 113

Letters to the Editor:

Electron Distribution in ZnO Crystals . . . . HUBERT M. JAMES AND VIVIAN A. JOHNSON 119

Super-Novae and the Neutron-Core Stars . . . . . FÉLIX CERNUSCHI 120

Tentative Theory of the Origin of Cosmic Rays . . . . . FÉLIX CERNUSCHI 120

Heat Capacity of Potassium Dihydrogen Phosphate at the Curie Point . . . . . C. C. STEPHENSON AND J. G. HOOLEY 121

Velocity of Sound in Liquid Helium under Pressure . . . . . J. C. FINDLAY, A. PITT, H. GRAYSON SMITH AND J. O. WILHELM 122

Multiple Scattering of Charged Particles . . . . . S. GOUDSMIT AND J. L. SAUNDERSON 122

Existence of Radioactive Al^{29} . . . . . W. J. HENDERSON AND R. L. DORAN 123

Transitions Between Glow and Arc Discharge . . . . . WALTER G. CADY 123

Initial Performance of the 60-Inch Cyclotron of the William H. Crocker Radiation Laboratory, University of California . . . . . ERNEST O. LAWRENCE, LUIS W. ALVAREZ, WILLIAM M. BROBECK, DONALD COOKSEY, DALE R. CORSON, EDWIN M. MCMILLAN, W. W. SALISBURY AND ROBERT L. THORNTON 124

LANCASTER PRESS, INC., LANCASTER, PA.